Planned Comparisons in ARM

Jun 9, 2023

Overview

- The term planned comparison is commonly used to describe hypothesis tests about parameters (e.g. treatment means) that are describe prior to collection of data.
- This term is used in contrast with unplanned or post-hoc hypothesis tests. Common multiple comparisons procedures, such as those used to produce mean separation letters (Fisher LSD, Turkey HSD, Student-Newman-Keuls, etc), are consider unplanned comparisons, primarily since they include all possible treatment comparisons.
- In ARM, planned comparisons can be specified in Protocol settings. This allows the designer of an experiment to specify treatment comparisons of scientific interest, when the treatment structure of an experiment is being decided.
- This is different than the procedure for specifying tests for post-hoc treatment test, which are selected during the reporting process.
- Contrasts are a general term for statistical tests on linear combinations of values derived from data. We will sometimes use this description as shorthand for linear combinations. `contrast` is used in SAS to specify a test involving user defined coefficients of linear combinations.
- In ARM we allow users to specify comparisons among different combinations of means. This allows testing of specific mean pairs or testing all pairs among a subset of treatments.

Overview, continued.

- Planned comparisons are usually specified as linear combinations of values derived from data. Sometimes, these linear combinations are referred to as contrasts.

- A hypothesis test stated as a linear combination can take the form, for t treatment means,

$$H_0 : \sum_{i=1}^{t} c_i \mu_i = a$$

- One constraint in the values of the user-defined coefficients c_i is that $\sum_{i=1}^{t} c_i = 0$

Overview, continued.

- A hypothesis of the form $H_0 : \mu_1 = \mu_2$ (that is mean of treatment 1 is equal to the mean of treatment 2) can be written as a linear combination of the form,

$$H_0 : \mu_1 - \mu_2 = 0$$

- with $c_1 = 1, c_2 = -1$ and $c_i = 0$ for $i \neq 1,2$.
- Clearly, $\sum_{i=1}^{t} c_i \mu_i = 1 + -1 + 0 + \ldots + 0 = 0.$

Overview, continued.

- Linear combinations are not limited to comparisons of single means. For example, we can specify that the average of treatments 1 and 2 be compared with the average of treatments 3, 4 and 5. The hypothesis would be written in the form

$$H_0 : (\mu_1 + \mu_2)/2 - (\mu_3 + \mu_4 + \mu_5)/3 = 0$$

- so that $\sum_{i=1}^{t} c_i \mu_i = \frac{1}{2} + \frac{1}{2} - \frac{1}{3} - \frac{1}{3} - \frac{1}{3} + 0 + \ldots + 0 = 0$
- The hypothesis could also be written as

$$H_0 : 3(\mu_1 + \mu_2) - 2(\mu_3 + \mu_4 + \mu_5) = 0$$

- and the constraint would be met. Internally, ARM uses the former convention for comparisons involving averages of means.

Linear Combination Statistic

- When hypotheses are stated as described above, the test statistic takes the form

$$\frac{\sum_{i=1}^{t} c_i \hat{\mu}_i - \sum_{i=1}^{t} c_i \mu_i}{\sqrt{\hat{\sigma}^2 \sum_{i=1}^{t} c_i^2 / n_i}}$$

  - where n_i is the number of observations for mean i and $\hat{\sigma}^2$ is the pooled error variance for the t means.

- This statistic is distributed as Student's t with $N - t$ degrees of freedom.

Linear Combination Statistic

- When there are only two means to be compared, say for treatments i and j, the formula

$$\frac{\sum_{i=1}^{t} c_i \hat{\mu}_i - \sum_{i=1}^{t} c_i \mu_i}{\sqrt{\hat{\sigma}^2 \sum_{i=1}^{t} c_i^2 / n_i}}$$

- reduces to

$$\frac{(1)\mu_i + (-1)\mu_j}{\sqrt{\frac{(1)^2 \hat{\sigma}^2}{n_i} + \frac{(-1)^2 \hat{\sigma}^2}{n_j}}}$$

which is the formula for a t-test of the difference of two means, with a pooled error term and (potentially) different number of replicates.

Example 1

- Table 1.1 from G. A. Milliken and D. E. Johnson. Analysis of Messy Data, Volume I Designed Experiments. Chapman and Hall/CRC, 2 Edition, 2009.
- This is entered as an ARM trial as `Milliken1.1.dat0`
- Milliken and Johnson illustrate inference on linear combinations using data from Table 1.1. In section 1.4, they propose the following hypothesis:
  A. Test $H_0 : \mu_3 = 30$
  B. Find a 95% confidence interval for μ_1
  C. Test $H_0 : \mu_4 = \mu_5$
  D. Test $H_0 : \mu_1 = (\mu_2 + \mu_3 + \mu_4)/3$
  E. Obtain a 90% confidence interval for $4\mu_1 - \mu_3 - \mu_4 - \mu_5 - \mu_6$

Example 1, SAS

- Milliken and Johnson provide SAS code that tests some of these hypothesis discussed in Chapter 1.

```
proc glm data=arm;
   class treatment;
   model assessment1 = treatment;
   lsmeans treatment / lines;
   estimate 'Ho: M4=M5' treatment 0 0 0 1 -1 0;
   estimate 'Ho: 3M1=M2+M3+M4' treatment 3 -1 -1 -1 0 0;
   estimate 'Ho: 3M1=M2+M3+M4_mn' treatment 3 -1 -1 -1 0 0/DIVISOR=3;
   estimate '4M1-M3-M4-M5-M6_mn' treatment 4 0 -1 -1 -1 -1/DIVISOR=4;
   contrast '4M1-M3-M4-M5-M6' treatment 4 0 -1 -1 -1 -1;
   contrast 'M4=M5 & 3M1=M2+M3+M4' treatment 0 0 0 1 -1 0,
                                  treatment 3 -1 -1 -1 0 0;
   contrast 'EQUAL MEANS 1' treatment 1 -1 0 0 0 0,
                            treatment 1 0 -1 0 0 0,
                            treatment 1 0 0 -1 0 0,
                            treatment 1 0 0 0 -1 0,
                            treatment 1 0 0 0 0 -1;
 run;
```

ARM Analysis : assessment1

Analysis of Variance

The GLM Procedure

Dependent Variable: assessment1

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
4M1-M3-M4-M5-M6	1	12.2044982	12.2044982	0.39	0.5320
M4=M5 & 3M1=M2+M3+M4	2	415.8586241	207.9293121	6.73	0.0023
EQUAL MEANS 1	5	694.4385987	138.8877197	4.49	0.0015

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Ho: M4=M5	8.50000000	2.38029755	3.57	0.0007
Ho: 3M1=M2+M3+M4	-9.11410256	5.49105204	-1.66	0.1020
Ho: 3M1=M2+M3+M4_mn	-3.03803419	1.83035068	-1.66	0.1020
4M1-M3-M4-M5-M6_mn	-1.10646853	1.76071732	-0.63	0.5320

Example 1 Planned Comparisons in ARM

- We don't typically report confidence intervals for means in ARM, so we won't consider hypotheses B and E. The remaining three hypothesis can be tested in ARM. We'll use the following terminology

- $H_{01} : \mu_3 = 30$ is a comparison of a single treatment mean against a constant. This is a **constant comparison**.

- $H_{02} : \mu_4 = \mu_5$ compares single mean against a single mean. This is a **paired comparison**.

- $H_{03} : \mu_1 = (\mu_2 + \mu_3 + \mu_4)/3$. This contrast compares a mean against the average of three other means. If either or both sides of the test equality is composed of more than one treatment, this is an **averaged comparison**.

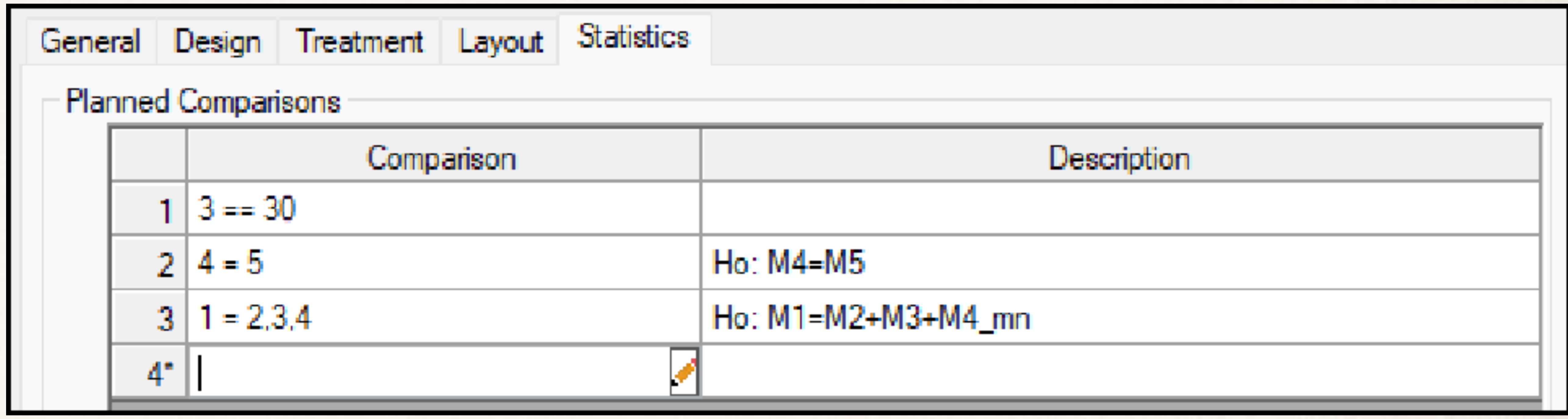
Example 1 Planned Comparisons in ARM

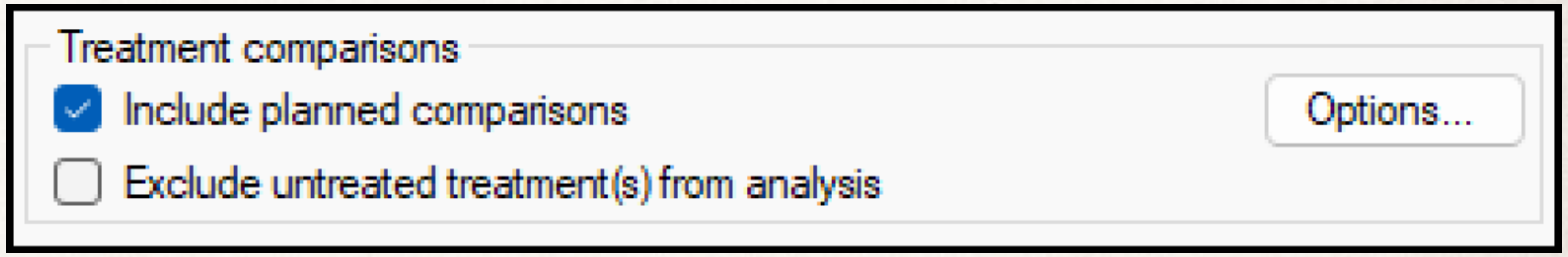
- Remember that the third hypothesis from the previous section can be specified using two equivalent linear combinations.
  - $H_{03} : \mu_1 = (\mu_2 + \mu_3 + \mu_4)/3$
  - $H_{03} : 3\mu_1 = \mu_2 + \mu_3 + \mu_4$
- Milliken and Johnson provide examples of both forms in their SAS code. The SAS output shows different values for the estimate of the linear combination, but both produce the same t statistics and p values.

Entering Comparisons in ARM

- Planned contrasts are entered in the Settings dialog, under the Statistics tab, or via the Report Options for AOV Means Table Report

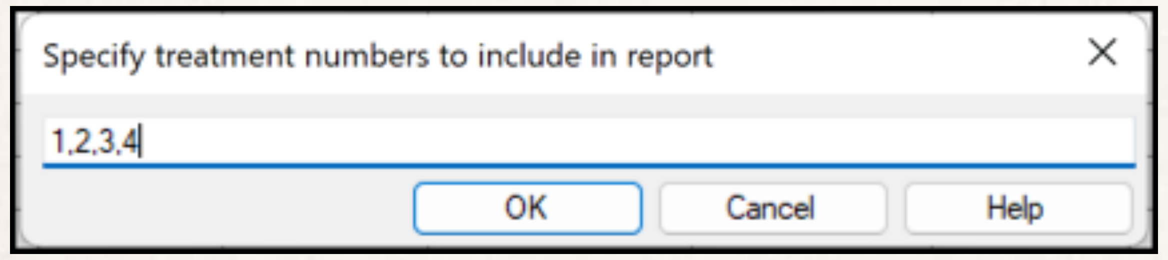
- Planned comparisons are included in the Settings dialog, so that they can be specified in protocols. This allows protocol writers to specify treatment comparisons of interest during trial design.



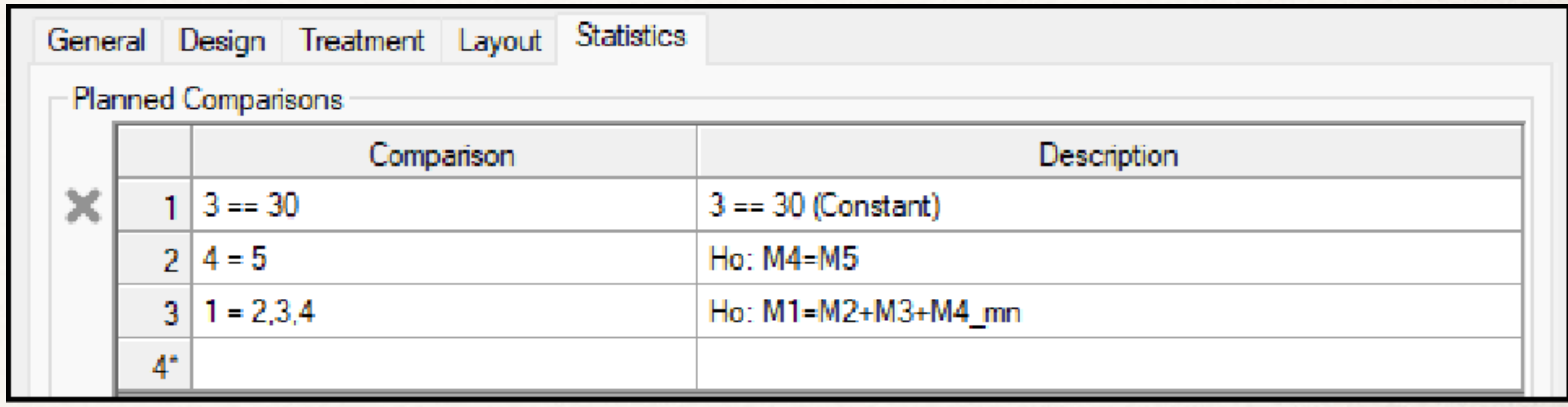


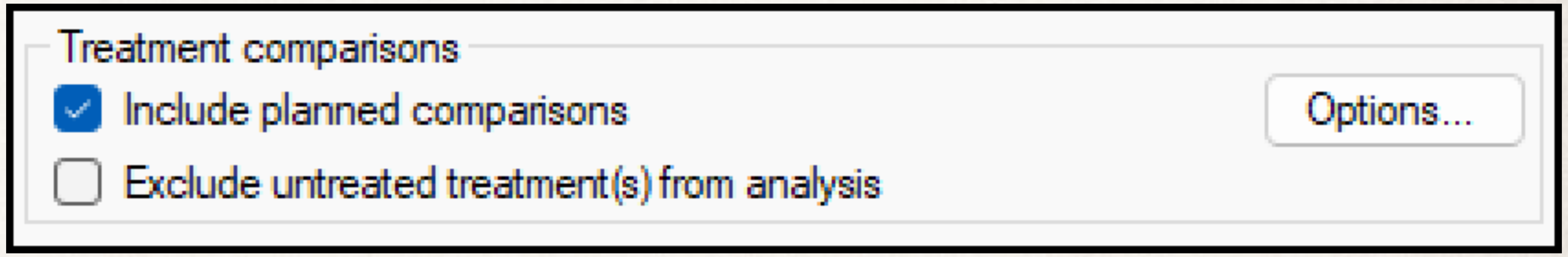
Entering Comparisons in ARM

- Contrast specifiers can be entered according to some simple rules:
    - Treatments to be compared are entered by treatment number.
    - A single equal sign (=) is used to separate treatments or groups of treatments. A double equal sign (==) is used to denote comparison of treatments against some user defined constant value.
    - Multiple treatments can be entered as space or comma-separated lists, with hyphens to denote a range of treatments. This is consistent the format use to select treatments for reports.



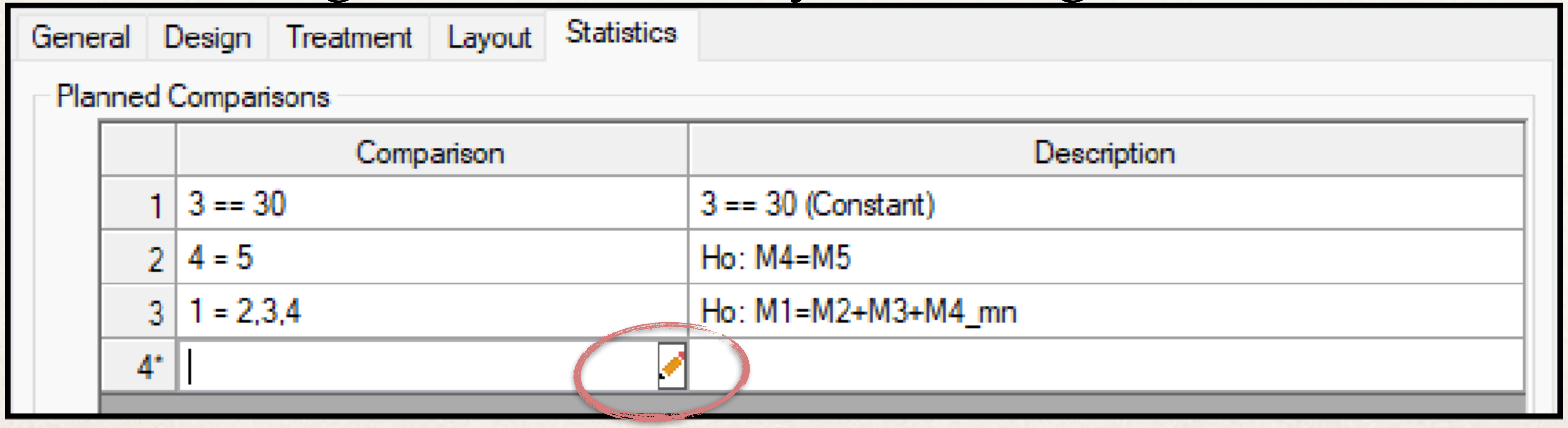
    - Multiple simultaneous tests are separated by semicolons (;)
- Descriptive text for report headings can be entered in the Description column. If no description is entered, a default description will be generated by ARM.





Comparison Wizard

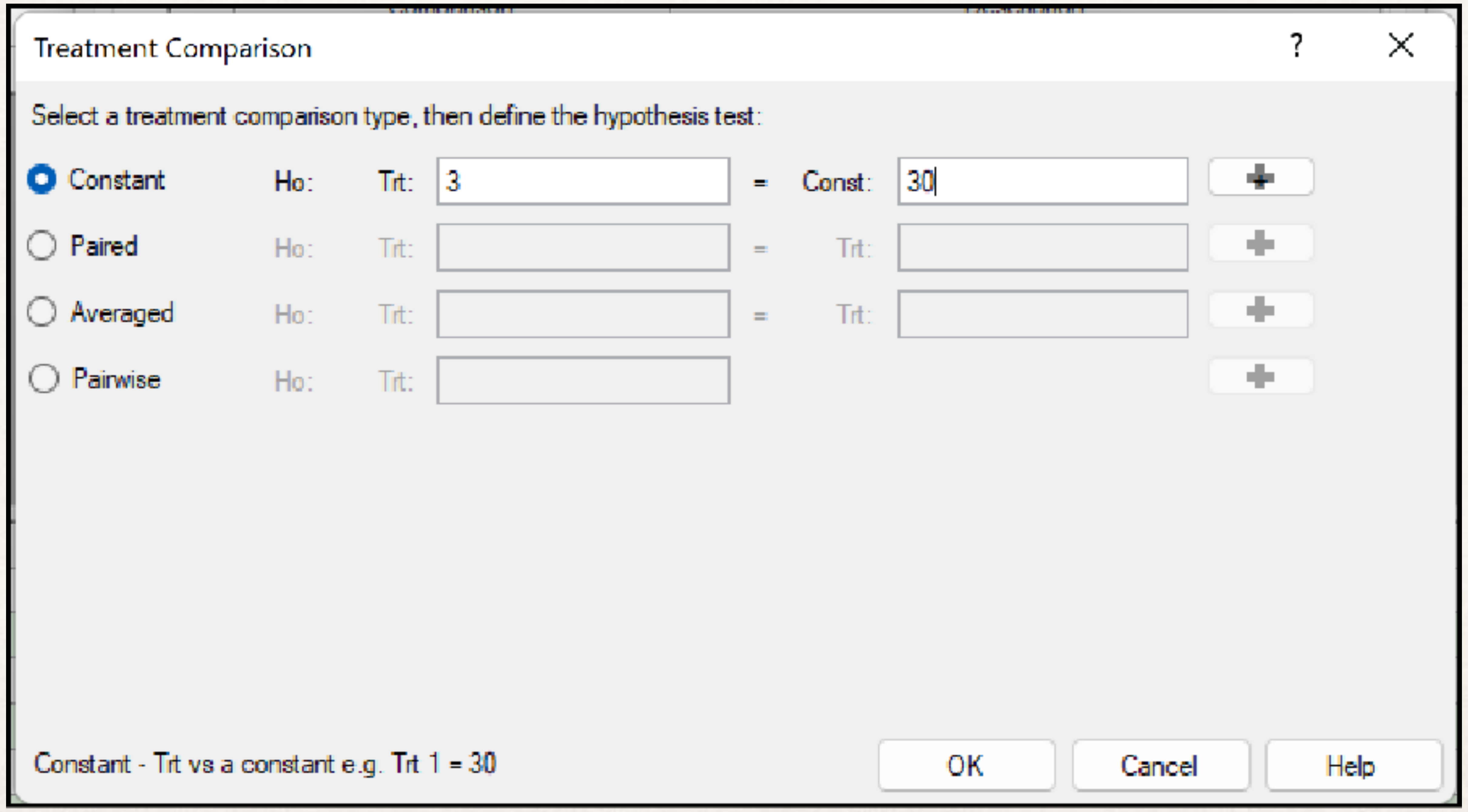
- While comparison specification text can be entered directly in the Planned Contrast table, ARM also provides a dialog to simplify entering comparison specification.
- This dialog is available by clicking the icon in the Comparison field:



	Comparison	Description
1	3 == 30	3 == 30 (Constant)
2	4 = 5	Ho: M4=M5
3	1 = 2,3,4	Ho: M1=M2+M3+M4_mn
4*		

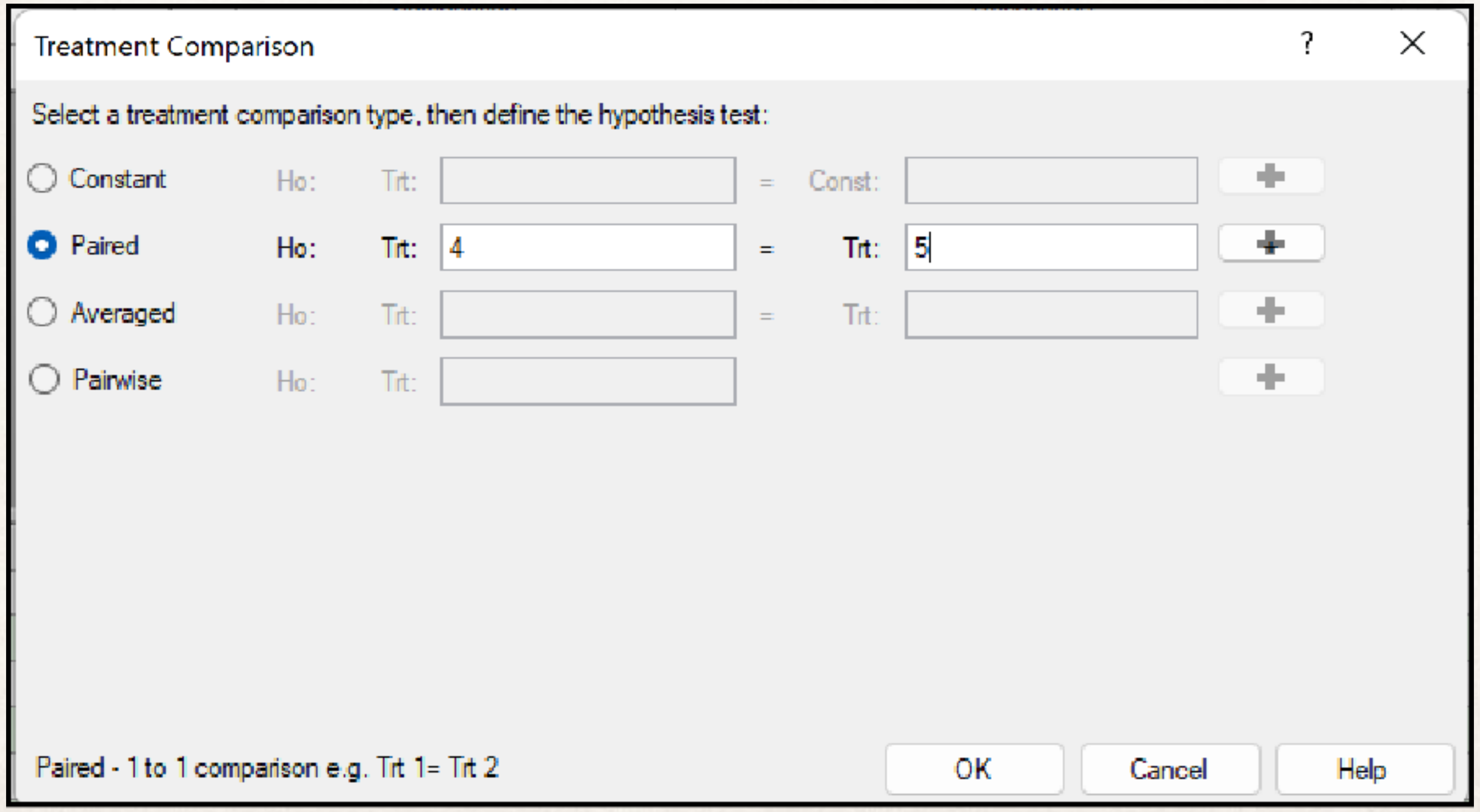
Comparison Wizard

- Dialog entries corresponding to
  - 3 == 30



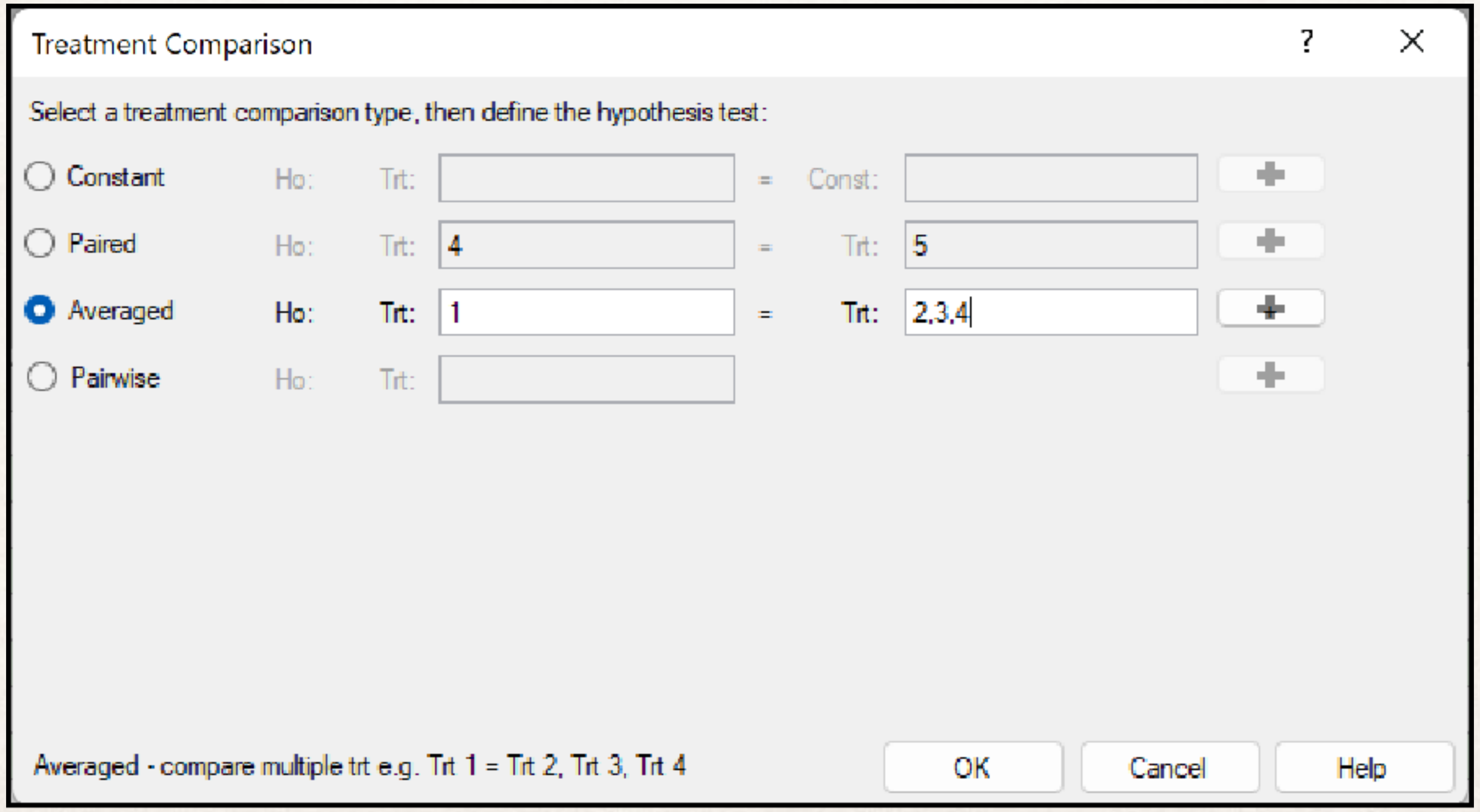
Comparison Wizard

- Dialog entries corresponding to
  - 4 == 5



Comparison Wizard

- Dialog entries corresponding to
  - 1 = 2,3,4



Comparisons in AOV Means Report

- Planned comparisons are included in a section below treatment means. The value of the contrast, calculated t or F statistic and associated p-value are reported for each contrast.
- Since the tests can be expressed as single equalities, a t-statistic is appropriate. The `Estimate` is the difference between the averages of two sides on both sides of the equality.
- The computed t-statistic and $P(>t)$ are also reported, where $P(>t)$ is the probability from a two-tailed t-test.

Character Rated Number of Decimals		Pulsation 6
Trt No.	Treatment Name	*
1	Task 1	31.923080 bc
2	Task 2	31.083336 bc
3	Task 3	35.800004 ab
4	Task 4	38.000004 a
5	Task 5	29.500003 c
6	Task 6	28.818185 c
Planned Comparisons 3 == 30 (Constant) Estimate t Value Pr > |t| Ho: M4=M5 Estimate t Value Pr > |t| Ho: M1=M2+M3+M4_mn Estimate t Value Pr > |t|		 5.8000000 3.2992670 0.0016093 8.5000000 3.5709821 0.0006942 -3.0380342 -1.6598099 0.1020029

Example 2 Simultaneous Comparisons

- Milliken and Johnson, continuing the examples of contrasts among treatment means from Table 1.1, propose two hypotheses to be simultaneously tested:

$$H_0 : \mu_4 - \mu_5 = 4 \text{ and } 3\mu_2 = \mu_2 + \mu_3 + \mu_4$$

- ARM does not currently support hypothesis of the form $\mu_4 - \mu_5 = 4$; we reserve the - character to denote a range of treatment numbers. Instead, we use the data from Exercise 1.2.
- These data are entered as the ARM trial `Milliken Ex 1.2.dat0`

Example 2 Simultaneous Comparisons

- In Exercise 1.2, Milliken and Johnson propose 8 parts; we can compute 5 in ARM:
  - 4) Use a t-statistic to test $H_0 : \mu_1 + \mu_2 - 2\mu_3 = 0$
  - 5) Use a F-statistic to test $H_0 : 2\mu_2 - \mu_4 - \mu_5 = 0$
  - 6) Use a t-statistic to test $H_0 : (\mu_1 + \mu_2 + \mu_3)/3 = (\mu_4 + \mu_5)/2$
  - 7) Use an F-statistic to test $H_0 : \mu_1 = \mu_2$ and $\mu_3 = \mu 4$
  - 8) Use an F-statistic to test $H_0 : \mu_1 + \mu_2 - 2\mu_3 = 0, 2\mu_2 - \mu_4 - \mu_5 = 0, (\mu_1 + \mu_2 + \mu_3)/3 = (\mu_4 + \mu_5)/2$
- Parts 4 and 6 are simple hypothesis tests that take the form of a t-statistic. In ARM, we compute t-statistic, not an F-statistic for part 5. Parts 7 and 8 contains multiple tests to be computed simultaneously. This requires, computationally, solving a system of equations. We'll briefly detail the process in the following section.

Mathematics for multiple comparisons.

- Suppose we have more than one linear combination, and we wish to test the significance of linear combinations taken simultaneously.
- Remember that a single linear combination takes the form

$$H_0 : \sum_{i=1}^{t} c_i \mu_i = a$$

- Several linear combinations can be written as a system of linear equations, of the form

$$\begin{aligned} c_{11}\mu_1 + c_{12}\mu_2 + \ldots + c_{1t}\mu_t &= a_1 \\ c_{21}\mu_1 + c_{22}\mu_2 + \ldots + c_{2t}\mu_t &= a_2 \\ &\vdots \\ c_{k1}\mu_1 + c_{k2}\mu_2 + \ldots + c_{kt}\mu_t &= a_k \end{aligned}$$

Mathematics for multiple comparisons.

- We can then write the hypothesis as

$$H_0 : \mathbf{C}\boldsymbol{\mu} = \mathbf{a}$$

- where

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1t} \\ c_{21} & c_{22} & \cdots & c_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ c_{k1} & c_{k2} & \cdots & c_{kt} \end{bmatrix}, \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_t \end{bmatrix}, \text{ and } \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix}$$

Mathematics for multiple comparisons.

- The sum of squares for testing $H_0 : \mathbf{C}\boldsymbol{\mu} = \mathbf{a}$ is given by

$$SS_{H0} = \left(\mathbf{C}\hat{\boldsymbol{\mu}} - \mathbf{a}\right)' \left(\mathbf{CDC}'\right)^{-1} \left(\mathbf{C}\hat{\boldsymbol{\mu}} - \mathbf{a}\right)$$

- where

$$\mathbf{D} = \begin{bmatrix} 1/n_1 & 0 & \dots & 0 \\ 0 & 1/n_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c_{k1} & c_{k2} & \dots & 1/n_t \end{bmatrix}$$

Mathematics for multiple comparisons.

- SS_{H0} has k (the number of rows in $\mathbf{C}$) degrees of freedom, so a mean square can be computed as SS_{H0}/k, and the F statistic for testing $H_0 : \mathbf{C}\boldsymbol{\mu} = \mathbf{a}$ is calculated by

$$F = \frac{SS_{H0}/k}{\hat{\sigma}^2}$$

Entering Multiple Comparisons

- In ARM, we enter multiple simultaneous comparisons as semi-colon separated statements.
- We enter the multiple tests as follows, using part 8 as an example:
  - $\mu_1 + \mu_2 - 2\mu_3 = 0$. This test is equivalent to $\mu_1 + \mu_2 = 2\mu_3$. We enter this as `1,2 = 3`; ARM will automatically determine the coefficients, so there is no need to enter the coefficient 2.
  - $2\mu_2 - \mu_4 - \mu_5 = 0$. As with the previous test, this is equivalent to $2\mu_2 = \mu_4 + \mu_5$.
  - $(\mu_1 + \mu_2 + \mu_3)/3 = (\mu_4 + \mu_5)/2$. This test can be entered in ARM as `1-3 = 4,5` or `1,2,3 = 4,5`. ARM internally computes the divisors 3 and 2.
  - Thus, the full simultaneous comparison is entered as

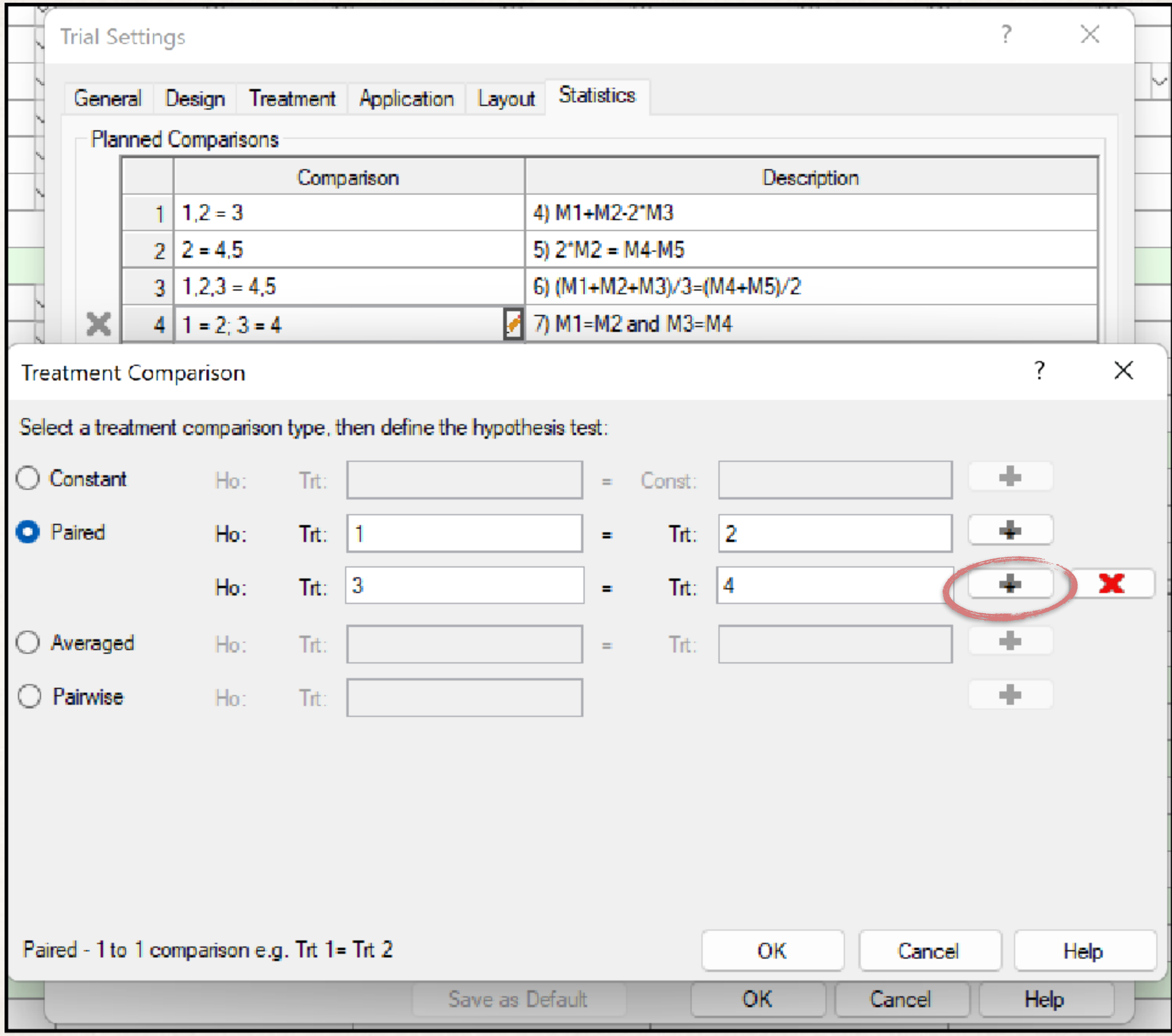
```
1,2 = 3; 2 = 4,5; 1-3 = 4,5
```

Planned Comparisons

	Comparison	Description
1	1,2 = 3	4) M1+M2-2*M3
2	2 = 4,5	5) 2*M2 = M4-M5
3	1,2,3 = 4,5	6) (M1+M2+M3)/3=(M4+M5)/2
4	1 = 2; 3 = 4	7) M1=M2 and M3=M4
5	1,2 = 3; 2 = 4,5; 1-3 = 4,5	8) (M1+M2)/2=M3,M2=(M4+M5)/2,(M1+M3+M4)/3-(M4+M5)/2
6*		

Entering Multiple Comparisons

- Multiple simultaneous comparisons can also be entered using the wizard dialog, by adding individual contrasts by selecting the  icon



Comparing ARM and SAS

- SAS code to reproduce the ARM output

```
title3 'Planned Comparisons';
proc glm data=arm;
  class treatment;
  model assessment1 = treatment;
  lsmeans treatment / lines;
  /* 1,2 = 3 */
  estimate '4) M1+M2-2*M3' treatment 1 1 -2 0 0/DIVISOR=2;
  /* 2 = 4,5 */
  estimate '5) 2*M2 = M4-M5' treatment 0 2 0 -1 -1/DIVISOR=2;
  /* 1,2,3 = 4,5 */
  estimate '6) (M1+M2+M3)/3=(M4+M5)/2' treatment 2 2 2 -3 -3/DIVISOR=6;
  /* 1=2;3=4 */
  contrast  '7) M1=M2 and M3=M4' treatment 1 -1 0 0 0,
                                 treatment 0 0 1 -1 0;
  /* 1,2=3;2=4,5;1,2,3=4,5 */
  contrast '8) (M1+M2)/2=M3,M2=(M4+M5)/2,(M1+M3+M4)/3-(M4+M5)/2'
           treatment 1 1 -2 0 0,
           treatment 0 2 0 -1 -1,
           treatment 2 2 2 -3 -3;
run;
```

Planned Comparisons

The GLM Procedure

Dependent Variable: assessment1

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
7) M1=M2 and M3=M4	2	5329.500000	2664.750000	5.57	0.0097
8) (M1+M2)/2=M3,M2=(M4+M5)/2,(M1+M3+M4)/3-(M4+M5)/2	3	6147.634703	2049.211568	4.28	0.0139

Parameter	Estimate	Standard Error	t Value	Pr > |t|
4) M1+M2-2*M3	-18.5555556	16.3035224	-1.14	0.2654
5) 2*M2 = M4-M5	30.8055556	11.2364274	2.74	0.0109
6) (M1+M2+M3)/3=(M4+M5)/2	22.1574074	10.5588998	2.10	0.0457

Comparing ARM and SAS

Planned Comparisons

The GLM Procedure

Dependent Variable: assessment1

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
7) M1=M2 and M3=M4	2	5329.500000	2664.750000	5.57	0.0097
8) (M1+M2)/2=M3,M2=(M4+M5)/2,(M1+M3+M4)/3-(M4+M5)/2	3	6147.634703	2049.211568	4.28	0.0139

Parameter	Estimate	Standard Error	t Value	Pr > |t|
4) M1+M2-2*M3	-18.5555556	16.3035224	-1.14	0.2654
5) 2*M2 = M4-M5	30.8055556	11.2364274	2.74	0.0109
6) (M1+M2+M3)/3=(M4+M5)/2	22.1574074	10.5588998	2.10	0.0457

Planned Comparisons	
4) M1+M2-2*M3	
Estimate	-18.56
t Value	-1.14
Pr > |t|	0.27
5) 2*M2 = M4-M5	
Estimate	30.81
t Value	2.74
Pr > |t|	0.01
6) (M1+M2+M3)/3=(M4+M5)/2	
Estimate	22.16
t Value	2.10
Pr > |t|	0.05
7) M1=M2 and M3=M4	
Mean square	2664.75
F value	5.57
Pr > F	0.01
8) (M1+M2)/2=M3,M2=(M4+M5)/2,(M1+M3+M4)/3-(M4+M5)/2	
Mean square	2049.21
F value	4.28
Pr > F	0.01

Example 3 Testing the Equality of All Means

- In Section 1.7 of "Analysis of Messy Data", Milliken and Johnson describe the contrasts to test the hypothesis

$$H_0 : \mu_1 = \mu_2 = \ldots \mu_t$$

- This is equivalent to simultaneously testing multiple hypothesis of the form

$$H_0 . \mu_1 - \mu_2 = 0 \text{ and } \mu_1 - \mu_3 = 0 \text{ and } \ldots \text{ and } \mu_1 - \mu_t = 0$$

- Other contrasts, consisting of linearly independent linear combinations can be constructed, but ARM uses this form when multiple pairwise contrasts among a set of means is specified.

Example 3 Linear Independence

- Suppose we have only three treatments to compare. Then

$$H_0 : \mu_1 = \mu_2 = \mu_3$$

- We could specify the contrast matrix as

$$\mathbf{C} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}$$

- However, this set of contrasts is not linearly independent. The final row can be written as a linear combination of the first two rows (i.e. row 2 - row 1). Thus, the correct contrast matrix would be

$$\mathbf{C} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$$

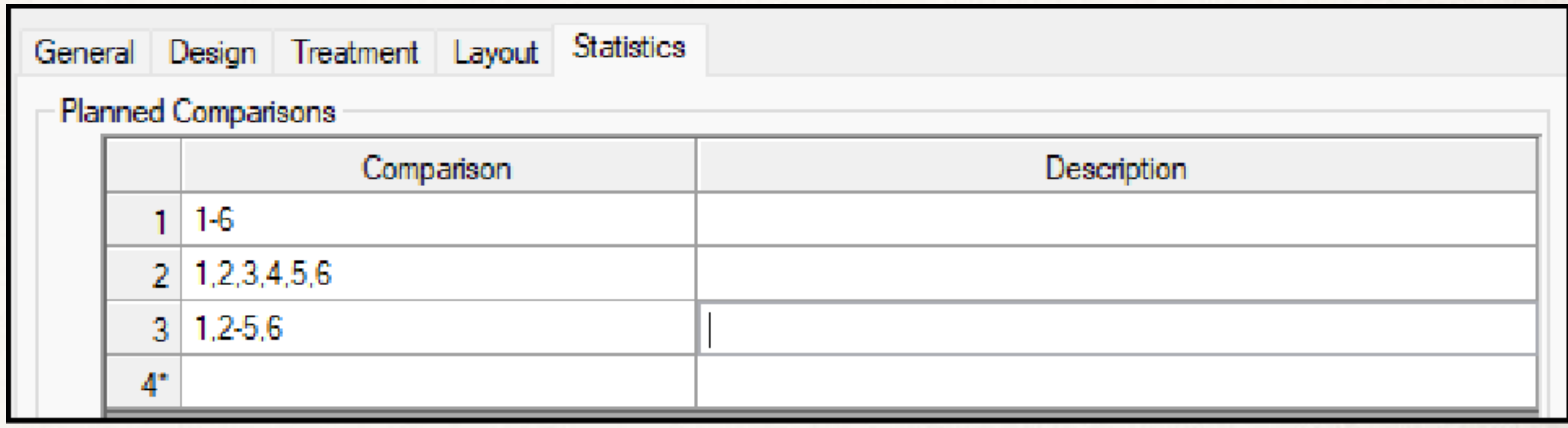
Example 3 Linear Independence

$$\mathbf{C} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$$

- ARM uses this form for comparisons of all means.
- Remember that the F test for treatment effect in an AOV table will have $t - 1$ degrees of freedom. Any linear combination involving multiple tests should not have more than $t - 1$ to be valid. Thus, we would be limited to $t - 1$ simultaneous comparisons.

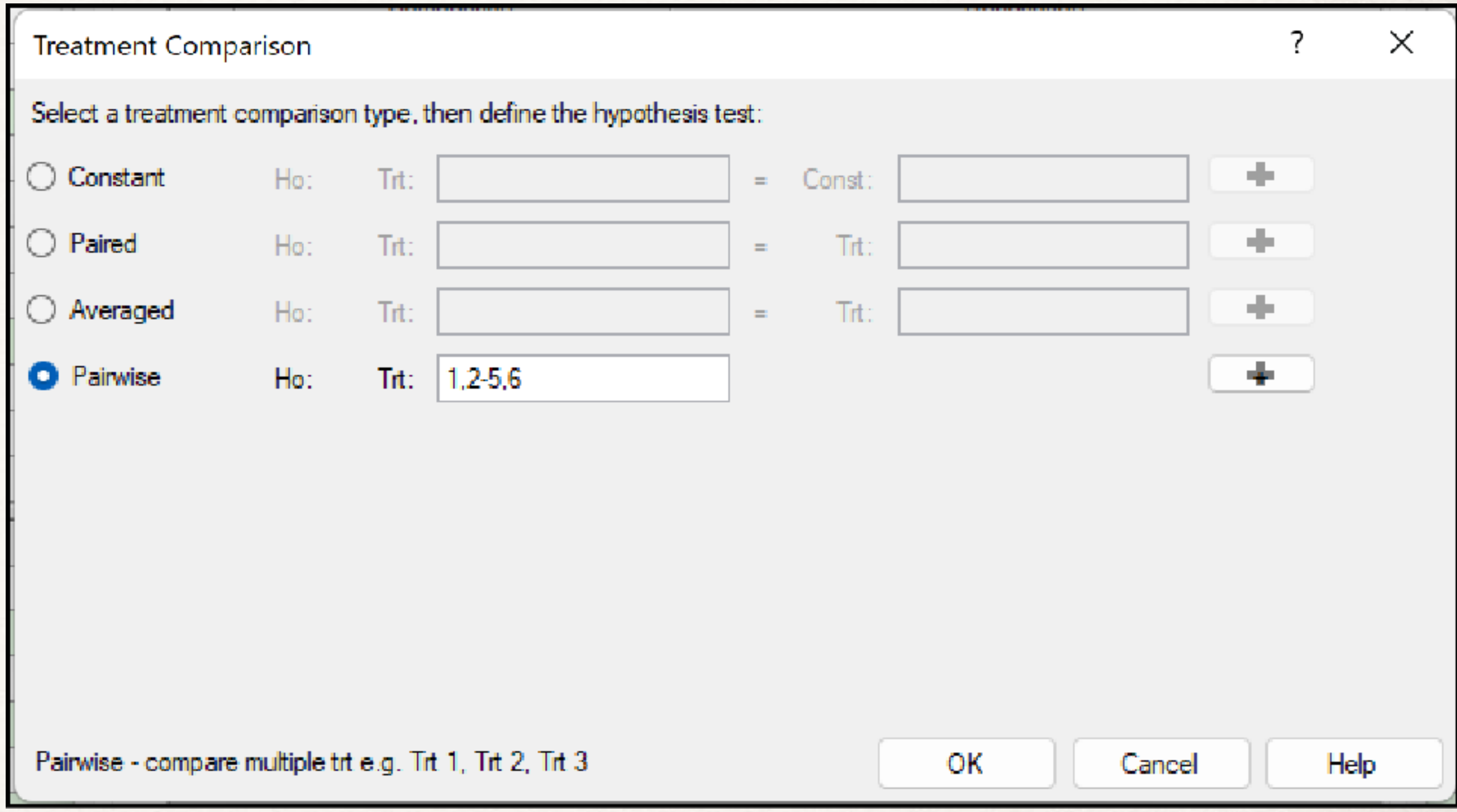
All-Pairwise Comparisons

- All pairwise contrasts among a set of means can be specified in ARM as a single list of treatment numbers, with no equal sign in the contrast specification.
- The treatment list can contain commas or hyphens.
- In the screenshot to the right, the three comparisons entered are equivalent.
- These comparisons are found in `Milliken1.1 Sec 7.dat0`



All-Pairwise Comparisons

- The Treatment Comparison wizard also allows these comparison specifications



All-Pairwise Comparisons

- All-pairwise comparisons is equivalent to the F test for treatments; that is, the F statistic tests a hypothesis of the form

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_t$$

- We see from the ARM report that these contrast specifications result in the same F ratios as the `Treatment F` in the AOV table. The reported `Contrast` value is the same as `Treatment Sum of Squares`.

Planned Comparisons	
1-6 (Pairwise)	
Mean square	138.8877197
F value	4.4941005
Pr > F	0.0014709
1,2,3,4,5,6 (Pairwise)	
Mean square	138.8877197
F value	4.4941005
Pr > F	0.0014709
1,2-5,6 (Pairwise)	
Mean square	138.8877197
F value	4.4941005
Pr > F	0.0014709

Completely Random (CRD) Least square estimation AOV For Pulsati

Source	DF	Sum of Squares	Mean Square	F	Prob(F)
Total	67	2610.514706			
Treatment	5	694.438599	138.88772	4.494	0.0015
Error	62	1916.076107	30.904453		

Example 4 Factorial Comparisons in One-way Treatment Structure

- I conducted an experiment to test six models of running shoe. The shoes were tested for running parameters - speed, stride length and stride rate. I tested each shoe six times, with a single training day as the experimental unit.

- I chose the shoes to represent two shoe brands (Nike and Brooks) and three relative weights (Lightweight Racing, Middleweight Tempo Trainer and Heavyweight Cushioned Trainer). This implies a factorial design (Brand x Weight). However, I can test the differences among Brands and among Shoe Weights using linear combinations.

Factor-based Comparisons

- See `6 Shoe Trial LTN.dat0`

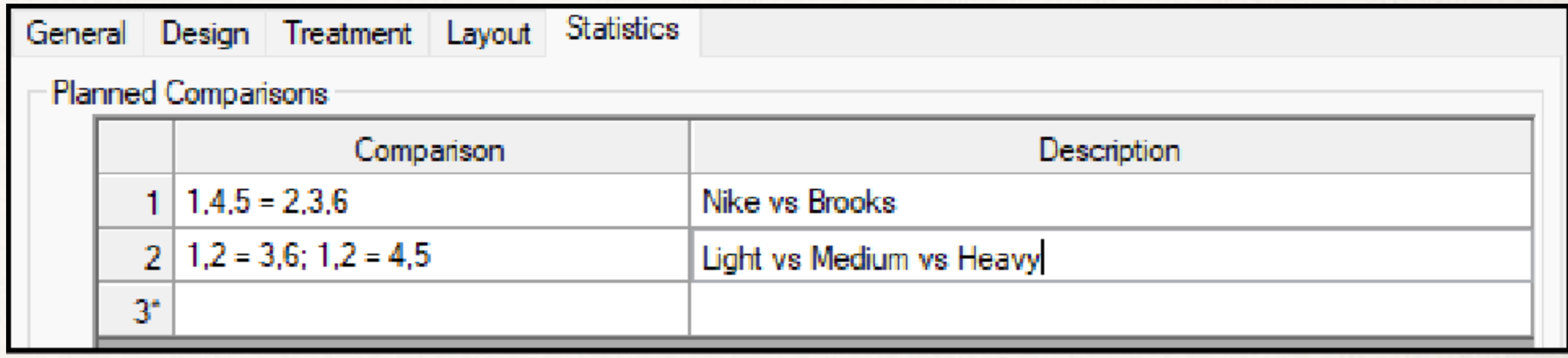
- In this trial, treatments 1,4 and 6 were Nike shoes, while treatments 2,3 and 5 were Brooks. I want to test the hypothesis that my running performance is better in Nike than Brooks. This implies a null hypothesis of the form

$$H_0 : \mu_1 + \mu_4 + \mu_6 = \mu_2 + \mu_3 + \mu_5$$

- With respect to weight, treatments 3 and 6 were light weight, treatments 4 and 5 were medium weight, and treatments 1 and 2 were heavy weight. I wish to test the hypothesis that my running performance is affected by shoe weight. This implies a composite null hypothesis of the forms

$$H_0 : \mu_1 + \mu_2 = \mu_4 + \mu_5 \text{ and } \mu_1 + \mu_2 = \mu_3 + \mu_6$$

- I don't need to test $\mu_1 + \mu_2 = \mu_3 + \mu_6$; that is implied if the preceding two tests are true.



General | Design | Treatment | Layout | Statistics

Planned Comparisons

	Comparison	Description
1	1,4,5 = 2,3,6	Nike vs Brooks
2	1,2 = 3,6; 1,2 = 4,5	Light vs Medium vs Heavy
3*		

Factor-based contrasts

- The first hypothesis

$$H_0 : \mu_1 + \mu_4 + \mu_6 = \mu_2 + \mu_3 + \mu_5$$

- is a single contrast, so is tested using a t statistic. This contrast test suggests a difference in running speed among shoe brands.

- The second hypothesis is composed of multiple comparisons.
  $H_0 : \mu_1 + \mu_2 = \mu_4 + \mu_5$ and $\mu_1 + \mu_2 = \mu_3 + \mu_6$
  This is tested using an F statistic. In this case, the test suggests differences in running speed among different weight classes.

Character Rated Rating Type Rating Unit/Min/Max ARM Action Codes Number of Decimals		Avg HR COUNT IID	Speed SPEED MPH, -, - T1 IID 2	Step Leng LENGTH FT, -, - T2 IID 2	Step Rate RATE PER SEC, -, - T3 IID 2
Trt No.	Treatment Name	*	*	*	*
1	LunarSwift Cushioned Trainer Nike	151.8 -	7.82 ab	4.02 -	171.23 -
2	Ghost Cushioned Trainer Brooks	147.8 -	7.64 b	3.94 -	171.22 -
3	Green Silence Racing Flat Brooks	151.3 -	7.88 ab	4.16 -	167.16 -
4	LunarFly Light Trainer Nike	153.3 -	7.83 ab	3.94 -	174.83 -
5	Launch Light Trainer Brooks	149.7 -	7.68 ab	3.92 -	172.45 -
6	Lunaracer2 Racing Flat Nike	153.0 -	7.92 a	4.21 -	168.32 -
Planned Comparisons					
Nike vs Brooks					
Estimate		3.11	0.123	0.048	1.185
t Value		2.35	2.572	0.507	0.418
Pr > |t|		0.03	0.018	0.618	0.681
Light vs Medium vs Heavy					
Mean square		16.77	0.095	0.205	99.970
F value		1.10	4.584	2.635	1.430
Pr > F		0.35	0.023	0.098	0.264

Contrasts versus Factorial AOV

- I intended this trial to be analyzable as factorial design. A factorial AOV suggests similar inferences about brand and weight effects as we would make with contrasts.
- We should note that contrasts do not test factorial (A x B) interactions.

Character Rated Rating Type Rating Unit/Min/Max ARM Action Codes Number of Decimals		Avg HR COUNT IID	Speed SPEED MPH, -, - T1 IID 2	Step Leng LENGTH FT, -, - T2 IID 2	Step Rate RATE PER SEC, -, - T3 IID 2
Trt No.	Treatment Name	*	*	*	*
1	LunarSwift Cushioned Trainer Nke	151.8 -	7.82 ab	4.02 -	171.23 -
2	Ghost Cushioned Trainer Brooks	147.8 -	7.64 b	3.94 -	171.22 -
3	Green Silence Racing Flat Brooks	151.3 -	7.88 ab	4.16 -	167.16 -
4	LunarFly Light Trainer Nike	153.3 -	7.83 ab	3.94 -	174.83 -
5	Launch Light Trainer Brooks	149.7 -	7.68 ab	3.92 -	172.45 -
6	Lunaracer2 Racing Flat Nike	153.0 -	7.92 a	4.21 -	168.32 -
Planned Comparisons					
Nike vs Brooks					
Estimate		3.11	0.123	0.048	1.185
t Value		2.35	2.572	0.507	0.418
Pr > |t|		0.03	0.018	0.618	0.681
Light vs Medium vs Heavy					
Mean square		16.77	0.095	0.205	99.970
F value		1.10	[illegible]	2.635	1.430
Pr > F		0.35	0.023	0.098	0.264

FACTORIAL/POOLED ERROR Least square estimation AOV For Speed SPEED

Source	DF	Sum of Squares	Mean Square	F	Prob(F)	HSD (.05)
Total	35	13.636019				
R	5	12.747728	2.549546	119.895	0.0001	
A	2	0.189188	0.094594	4.448	0.0222	0.15
B	1	0.136563	0.136563	6.422	0.0179	0.10
AB	2	0.030921	0.015460	0.727	[illegible]	0.26
ERROR	25	0.531620	0.021265			

Example 4 Not all treatments are equally interesting

- I performed a second shoe trial, with the same basic design (a 6x6 Latin Square) as the first. However, I was less systematic in selecting shoes to enter in the trial. The trial was entered as `Shoe Technology.dat0`.

- I included two models (Nike Pegasus and Adidas Marathon) that best represent "typical" running shoes. Two models were Karhu brand, and include a mid-foot fulcrum that's supposed to speed the heel-to-toe transition. One model was Newton brand, with forefoot lugs that are suppose to speed to toe-off in stride.

- The sixth model was Nike Zoom Elite, which I didn't have much particular interest in testing, but I needed to have six shoe models to fill out a 6x6 Latin Square.

Example 4, continued.

- Since I'm not particularly interesting in the Zoom Elite (treatment 6), I could just perform an analysis of treatments 1-5. However, that would invalidate the planned design, so, as best practice, I should analyze all treatments.
- I can use contrasts to make the specific comparisons I'm most interesting in testing:
  - Are the two Karhu models (treatments 1 and 2) different from the traditional running shoes (treatments 4 and 5)?
  - Is the Newton model (treatment 3) different from the traditional running shoes (treatments 4 and 5)?

Example 4, continued.

- There are other statistical arguments for including a treatment (6, Zoom Elite) in the analysis:
    - Including all treatments provide a better estimate of error. When we use the AOV Residual MS as an error term, we are effectively pooling the standard deviation of each treatment. If we exclude treatment 6, we would be pooling 5 standard deviations instead of 6; this uses 5/6 of the information we have from the experiment.
    - If we exclude treatments, we will have fewer degrees of freedom for error. This will increase the magnitude of the critical value for means tests. For example, a 6 treatment RCB trial of 6 replicates has an error d.f. of 24; removing a treatment reduces error d.f. to 19. This increases a critical *t* value from 1.71 to 1.73.

Example 4 Comparisons

- There are three comparisons of interest:
- 4,5 = 1,2 (Traditional vs Fulcrum)
  - This specifies the simple comparison of the means of treatments 4 and 5 (traditional shoes) against the means of treatments 1 and 2 (fulcrum shoes). The traditional shoes averaged slightly faster (0.032 m/s) but this is not significant ($P(>t)$=0.297).

		Cadence	Speed	Stride Leng
Character Rated Rating Type Rating Unit/Min/Max Data Entry Date ARM Action Codes Number of Subsamples Number of Decimals		14 2	SPEED m/s, -, - T1 14 2	LENGTH m, -, - T2 2
Trt No.	Treatment Name	*	*	*
1	Fast 2	89.55 -	3.36 -	2.25 -
2	Flow	90.45 -	3.38 -	2.24 -
3	Gravity	89.71 -	3.40 -	2.27 -
4	Marathon 10	89.85 -	3.43 -	2.29 -
5	Pegasus 28	89.85 -	3.38 -	2.25 -
6	Zoom Elite	90.03 -	3.36 -	2.23 -
Planned Comparisons				
Traditional vs Fulcrum				
Estimate		-0.151	0.032	0.026
t Value		-0.578	1.072	1.632
Pr > |t|		0.569	0.297	0.118
Traditional vs Newton				
Estimate		0.136	0.005	0.000
t Value		0.424	0.148	-0.003
Pr > |t|		0.676	0.884	0.998
Excluding Zoom Elite				
Mean square		0.698	0.004	0.002
F value		1.706	0.735	1.372
Pr > F		0.188	0.579	0.279

Example 4 Contrasts

- 4,5 = 3 (Traditional vs Newton)
    - This specifies the simple comparison of the means of treatments 4 and 5 (traditional shoes) against the mean of treatments 3 (forefoot lugs). The lugged shoe was slightly slower (0.005 m/s) but this difference was not significant ($P(>t) = 0.884$.

Character Rated		Cadence	Speed	Stride Leng
Rating Type			SPEED	LENGTH
Rating Unit/Min/Max			m/s, -, -	m, -, -
Data Entry Date				
ARM Action Codes			T1	T2
Number of Subsamples		14	14	
Number of Decimals		2	2	2
Trt No.	Treatment Name	*	*	*
1	Fast 2	89.55 -	3.36 -	2.25 -
2	Flow	90.45 -	3.38 -	2.24 -
3	Gravity	89.71 -	3.40 -	2.27 -
4	Marathon 10	89.85 -	3.43 -	2.29 -
5	Pegasus 28	89.85 -	3.38 -	2.25 -
6	Zoom Elite	90.03 -	3.36 -	2.23 -
Planned Comparisons				
Traditional vs Fulcrum				
Estimate		-0.151	0.032	0.026
t Value		-0.578	1.072	1.632
Pr > |t|		0.569	0.297	0.118
Traditional vs Newton				
Estimate		0.136	0.005	0.000
t Value		0.424	0.148	-0.003
Pr > |t|		0.676	0.884	0.998
Excluding Zoom Elite				
Mean square		0.698	0.004	0.002
F value		1.706	0.735	1.372
Pr > F		0.188	0.579	0.279

Example 4 Contrasts

- 1-5 (All shoes excluding Zoom Elite)
- This specifies all pair-wise comparisons among treatments 1-5, excluding the model I don't care about. The contrast value is a simple sum of squares, and is not interpretable in the same units as the means.
  - This contrast requires an *F* test, since there is no single comparison. The *F* ratio (0.735) might* be comparable to the *F* value obtained by analyzing only treatments 1-5 in a standard AOV.
    - * This trial as implemented as a Latin square, so excluding a treatment is not practical.

Character Rated		Cadence	Speed	Stride Leng
Rating Type			SPEED	LENGTH
Rating Unit/Min/Max			m/s, -, -	m, -, -
Data Entry Date				
ARM Action Codes			T1	T2
Number of Subsamples		14	14	
Number of Decimals		2	2	2
Trt No.	Treatment Name	*	*	*
1	Fast 2	89.55 -	3.36 -	2.25 -
2	Flow	90.45 -	3.38 -	2.24 -
3	Gravity	89.71 -	3.40 -	2.27 -
4	Marathon 10	89.85 -	3.43 -	2.29 -
5	Pegasus 28	89.85 -	3.38 -	2.25 -
6	Zoom Elite	90.03 -	3.36 -	2.23 -
Planned Comparisons				
Traditional vs Fulcrum				
Estimate		-0.151	0.032	0.026
t Value		-0.578	1.072	1.632
Pr > |t|		0.569	0.297	0.118
Traditional vs Newton				
Estimate		0.136	0.005	0.000
t Value		0.424	0.148	-0.003
Pr > |t|		0.676	0.884	0.998
Excluding Zoom Elite				
Mean square		0.698	0.004	0.002
F value		1.706	0.735	1.372
Pr > F		0.188	0.579	0.279

Example 5

- Table 2.3, Gomez and Gomez. Statistical Procedures for Agricultural Research. John Wiley and Sons, 2 Edition, 1984. These data were entered as `Gomez 2.3.dat0`
- The treatments include various combinations of types, rates and application timings of postemergence herbicides. The treatment structure suggests some combinations of particular interest.
- I've defined set of planned treatment comparisons, plus added an overall F-test of the four planned comparisons. The F-test can be used as a validation of the four planned comparisons, when made simultaneously. This is comparable to the report setting "Only when significant AOV treatment P(F)"

Example 5 - Planned Comparisons

Treatments - Line 2

Trt Line	Trt No.	Type	Treatment Name	For m	For m	For m	De	Rate	Rate Unit	Ot he	Ot her	Appl Timing	A p	Appl Descriptio
1	1	HERB	Propanil					2.0	kg AI/ha			POEMCR		21 DAS
2	1	HERB	Bromoxynil					0.25	g/100 kg			POEMCR		
3	2	HERB	Propanil					3.0	kg AI/ha			POEMCR		28 DAS
4	2	HERB	2,4-D-B					1.0	kg AI/ha			POEMCR		
5	3	HERB	Propanil					2.0	kg AI/ha			POEMCR		14 DAS
6	3	HERB	Bromoxynil					0.25	kg AI/ha			POEMCR		
7	4	HERB	Propanil					2.0	kg AI/ha			POEMCR		14 DAS
8	4	HERB	Ioxynil					0.5	kg AI/ha			POEMCR		
9	5	HERB	Propanil					3.0	kg AI/ha			POEMCR		21 DAS
10	5	HERB	CHCH					1.50	kg AI/ha			POEMCR		
11	6	HERB	Phenyedipham					1.5	kg AI/ha			POEMCR		14 DAS
12	7	HERB	Propanil					2.0	kg AI/ha			POEMCR		28 DAS
13	7	HERB	Bromoxynil					2.5	kg AI/ha			POEMCR		
14	8	HERB	Propanil					3.0	kg AI/ha			POEMCR		28 DAS
15	8	HERB	2,4-D-IPE					1.0	kg AI/ha			POEMCR		
16	9	HERB	Propanil					2.0	kg AI/ha			POEMCR		28 DAS
17	9	HERB	Ioxynil					0.5	kg AI/ha			POEMCR		
18	10	CULT	Handweeded											15 and 35 D
19	11	CHK	Control											

Planned Comparisons

	Comparison	Description
1	1 = 3; 4 = 9; 1 = 7; 6 = 10	Omnibus test
2	1 = 3	Bromoxnil at 21 DAS vs Bromoxnil at 14 DAS
3	4 = 9	Ioxynil at 14 DAS vs Ioxynil at 28 DAS
4	1 = 7	Bromoxnil (0.25 kg AI/ha) at 21 DAS vs Bromoxnil (2.5 kg AI/ha)
5	6 = 10	Phenydiphem vs Handweeded
6*		

- Omnibus test - all planned comparisons, simultaneously.
- Treatment 1 (Propanil and Bromoxnil at 21 DAS) versus Treatment 3 (Propanil and Bromoxnil at 14 DAS)
- Treatment 4 (Propanil and Ioxynil at 14 DAS) vs Treatment 9 (Propanil and Ioxynil at 28 DAS)
- Treatment 1 (Propanil and Bromoxnil (0.25 kg AI/ha) at 21 DAS) versus Treatment 7 (Propanil and Bromoxnil (2.5 kg AI/ha) at 28 DAS)
- Treatment 6 (Phenydiphem) vs Treatment 10 (Handweeded)

Example 5

- Omnibus test
  - Similar to Fisher's protected LSD, this test suggests that at least one of the planned comparisons is significant.

Planned Comparisons	
Omnibus test	
Mean square	1082753.08
F value	6.13
Pr > F	<0.01
Bromoxnil at 21 DAS vs Bromoxnil at 14 DAS	
Estimate	695.75
t Value	2.34
Pr > |t|	0.03
Ioxynil at 14 DAS vs Ioxynil at 28 DAS	
Estimate	868.50
t Value	2.92
Pr > |t|	<0.01
Bromoxnil (0.25 kg AI/ha) at 21 DAS vs Bromoxnil (2.5 kg AI/ha) at 21 DAS	
Estimate	1160.50
t Value	3.91
Pr > |t|	<0.01
Phenydiphem vs Handweeded	
Estimate	-233.33
t Value	-0.73
Pr > |t|	0.47

Example 5

- 1 = 3 (Bromoxnil at 21 DAS vs Bromoxnil at 14 DAS)
  - Letters for a mean comparison test using Student–Newman–Keuls (SNK) are *a* and *ab*. When allowing for multiple comparisons, we would not assert that treatments 1 and 3 are different. The single user contrast, on the other hand, suggests the treatments are indeed different.

Planned Comparisons	
Omnibus test	
Mean square	1082753.08
F value	6.13
Pr > F	<0.01
Bromoxnil at 21 DAS vs Bromoxnil at 14 DAS	
Estimate	695.75
t Value	2.34
Pr > |t|	0.03
Ioxynil at 14 DAS vs Ioxynil at 28 DAS	
Estimate	868.50
t Value	2.92
Pr > |t|	<0.01
Bromoxnil (0.25 kg AI/ha) at 21 DAS vs Bromoxnil (2.5 kg AI/ha) at 21 DAS	
Estimate	1160.50
t Value	3.91
Pr > |t|	<0.01
Phenydiphem vs Handweeded	
Estimate	-233.33
t Value	-0.73
Pr > |t|	0.47

Example 5

- 4 = 9 (Ioxynil at 14 DAS vs Ioxynil at 28 DAS)
  - Similarly, this treatment pair has overlapping letters (*ab* and *b*), but the single user contrast suggests a difference.

Planned Comparisons	
Omnibus test	
Mean square	1082753.08
F value	6.13
Pr > F	<0.01
Bromoxnil at 21 DAS vs Bromoxnil at 14 DAS	
Estimate	695.75
t Value	2.34
Pr > |t|	0.03
Ioxynil at 14 DAS vs Ioxynil at 28 DAS	
Estimate	868.50
t Value	2.92
Pr > |t|	<0.01
Bromoxnil (0.25 kg AI/ha) at 21 DAS vs Bromoxnil (2.5 kg AI/ha) at 21 DAS	
Estimate	1160.50
t Value	3.91
Pr > |t|	<0.01
Phenydiphem vs Handweeded	
Estimate	-233.33
t Value	-0.73
Pr > |t|	0.47

Example 5

- 1 = 7 (Bromoxnil (0.25 kg AI/ha) at 21 DAS vs (Bromoxnil (2.5 kg AI/ha) at 28 DAS)
  - This treatment pair has been different letters using SNK. This is consistent with the planned user contrast. However, this pair has both different rates and application dates, so we can't be certain if one or both cause the measured difference

Planned Comparisons	
Omnibus test	
Mean square	1082753.08
F value	6.13
Pr > F	<0.01
Bromoxnil at 21 DAS vs Bromoxnil at 14 DAS	
Estimate	695.75
t Value	2.34
Pr > |t|	0.03
Ioxynil at 14 DAS vs Ioxynil at 28 DAS	
Estimate	868.50
t Value	2.92
Pr > |t|	<0.01
Bromoxnil (0.25 kg AI/ha) at 21 DAS vs Bromoxnil (2.5 kg AI/ha) at 21 DAS	
Estimate	1160.50
t Value	3.91
Pr > |t|	<0.01
Phenydiphem vs Handweeded	
Estimate	-233.33
t Value	-0.73
Pr > |t|	0.47

Example 5

- 6=10
  - We cannot determine a statistical difference between Phenydiphem and hand-weeding. When using LSD for multiple comparisons, the two treatments are assigned *bcd* and *bc*, respectively.

Planned Comparisons	
Omnibus test	
Mean square	1082753.08
F value	6.13
Pr > F	0.00
Bromoxnil at 21 DAS vs Bromoxnil at 14 DAS	
Estimate	695.75
t Value	2.34
Pr > |t|	0.03
Ioxynil at 14 DAS vs Ioxynil at 28 DAS	
Estimate	868.50
t Value	2.92
Pr > |t|	0.01
Bromoxnil (0.25 kg AI/ha) at 21 DAS vs Bromoxnil (2.5 kg AI/ha) at 21 DAS	
Estimate	1160.50
t Value	3.91
Pr > |t|	0.00
Phenydiphem vs Handweeded	
Estimate	-233.33
t Value	-0.73
Pr > |t|	0.47

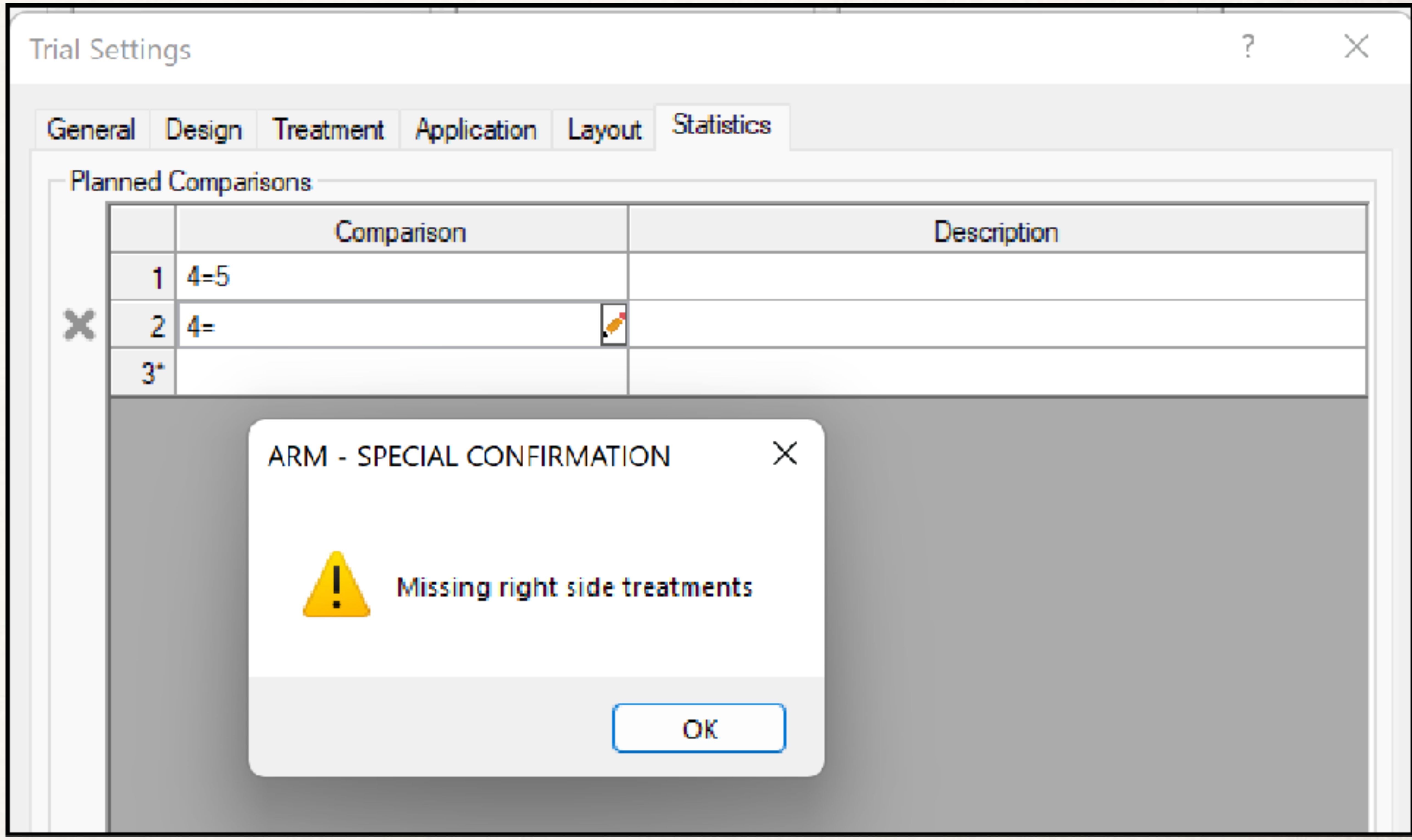
Example 6
Transformations

- We've used `Milliken2.1.dat0` as a test case for comparing IID, AL and AR analysis.
- When columns specific AL, AS or AA, the planned comparisons are applied to the transformed means and standard errors.
- Planned comparisons are not compatible with AR transformation; the rank-based analysis of AR is not compatible with the mathematics behind planned comparisons.

Character Rated ARM Action Codes		Score IID	Score AL	Score AR
Trt No.	Treatment Name	*	* AL	*
1	No Drug	4.6 c	0.6 b	5.7 d
2	Drug 1	11.7 a	1.1 a	18.0 b
3	Drug 2	8.6 b	1.0 a	11.4 c
4	Drugs 1 and 2	13.8 a	1.2 a	24.4 a
Planned Comparisons				
1 = 2 (Paired)				
Estimate		-7.10	-0.49	.
t Value		-4.57	-4.05	.
Pr > |t|		<0.01	<0.01	.
1 = 3 (Paired)				
Estimate		-4.05	-0.35	.
t Value		-2.81	-3.12	.
Pr > |t|		<0.01	<0.01	.
2 = 4 (Paired)				
Estimate		-2.08	-0.07	.
t Value		-1.38	-0.56	.
Pr > |t|		0.18	0.58	.
3 = 4 (Paired)				
Estimate		-5.12	-0.21	.
t Value		-3.68	-1.88	.
Pr > |t|		<0.01	0.07	.
Linear Independent				
Mean square		115.93	0.44	.
F value		14.91	9.26	.
Pr > F		<0.01	<0.01	.
Linear Dependent				
Mean square		57.96	0.22	.
F value		7.45	4.63	.
Pr > F		<0.01	<0.01	.

Specification Errors

- When the OK button is selected from the User Comparisons dialog, each comparison is scanned and the first error found is reported.
- In the following slides, we show different types of contrast specification errors and their associated error messages.



Comparison Errors

- We’ve taken care to define a method of specifying comparisons that simplifies using linear combinations.
- However, it will be possible to enter a comparison specification that cannot be interpreted by ARM. Some errors can be found when the comparison are specified; other errors won’t be discovered until comparison are computed.
- If an error is discovered during computing comparison, there will be missing values in the contrast table. ARM will display these error conditions in the report message screen.
- When an error can be detected when the comparison is entered, the comparison text will be displayed in red italicized text, and the corresponding table item will have a tooltip briefly describing the error.
- In the next section, we will outline the kinds of errors ARM can detect.

Example 1 Error Messages

- 3 == 30
  - ***a*** *== 30, 3 ==* ***a***
    - Non-numeric values cannot be entered is the treatment list fields.
  - ***30*** *== 30*
    - Invalid treatment number 30. The number on the LHS is not found in the treatment list.
  - ***3*** *=* ***3***
    - Duplicate treatments
  - *30 ==* ***3-10***
    - Right hand side must be a numeric value
  - *3* **= =** *30*
    - Extra =

Example 1 Error Messages

- $4 = 5$
  - $4 =$
    - Missing right side treatments
  - $= 5$
    - Missing left side treatments

Example 1 Error Messages

- 1 = 2,3,4
  - *1 = 2,3,4,*
    - Missing right-side treatments
  - *1 = 2 = 3,4*
    - Extra =

- 1 = 2-4
  - *1 = 2-*
    - Invalid treatment range
  - *1 = 2-3-4*
    - Invalid treatment range

Example 2 Error Messages

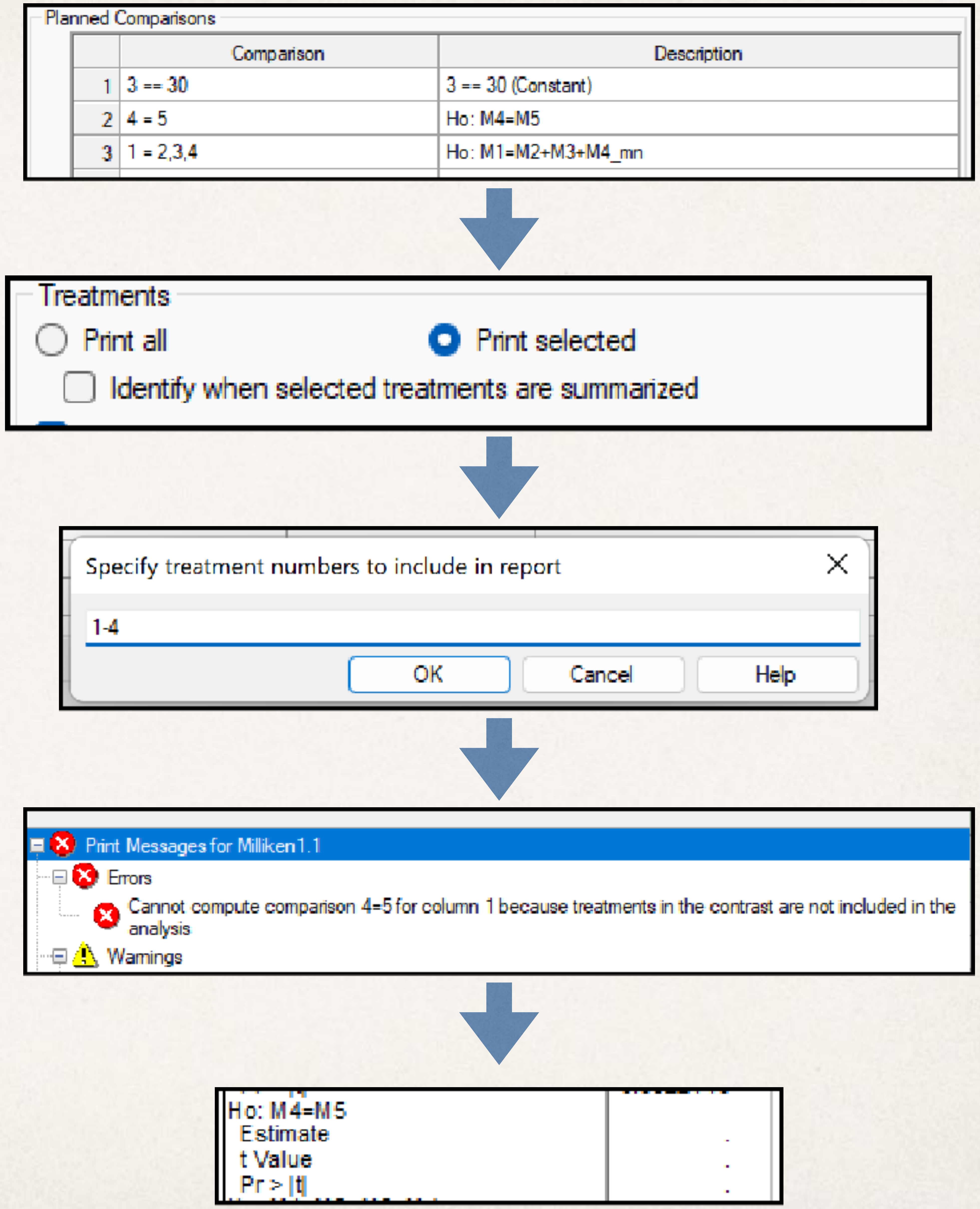
- 1,2 = 3; 2 = 4,5; 1-3 = 4,5
  - *1,2 =* ***3*** ***2*** *= 4,5 1-3 = 4,5*
    - Extra =
  - ***1 2*** *= 3; 2 = 4,5; 1-3 = 4,5*
    - Non-numeric treatment value
  - *1,2 = 3; 2 = 4,5;* ***1-*** *= 4,5*
    - Invalid treatment range

Example 3 Error Messages

- 1-6
  - *1-*
    - Invalid treatment range
- 1,2-5,6
  - *1 2-5,6*
    - Non-numeric treatment value
  - *1,2-,6*
    - Invalid treatment range

Analysis Errors

- Some errors can't be detected when the contrast specification is entered.
- For example, the specified treatments may not be included in the analysis (for example, when the "Print selected" Treatments option is chosen).
- These types of will result in missing values in the report.



Analysis Errors

- Types of analysis errors include
  - Treatments in specification not included in analysis
  - Standard error or error variance is 0
  - Missing treatment means
  - Contrasts incompatible with action codes

Treatments in specification not included in analysis

- This type of error can arise under two common circumstances:
  - Treatment number in contrast specification does not exist in treatment list
  - Treatment number in contrast specification not included in analysis.
    - The “Print selected” option in general is not compatible with planned comparisons. Sometimes this option may be used to “get rid” of treatments that are not of scientific interest. User contrasts can be used to the achieve the same effect, but have two benefits:
      - All plots are included in the analysis, so there is more information available in the analysis w.r.t spatial variability (there are no “missing plots”)
      - There are more degrees of freedom for error, so comparisons can be made with more statistical precision

Standard error or error variance is 0

- As with mean separation letters, there must an error term to compute critical values (i.e. LSD, HSD). If error is effectively 0, there is no usable error term and we can provide no hypothesis tests.

Missing treatment means

- In the course of an experiment, plot assessments for specific treatments may be lost or not measured. In such a case, the treatment mean is not available, so the contrast value cannot be computed.

Summary

- Linear combinations are a tool to test specific hypothesis concerning treatment means. Linear combinations can be used when omnibus tests (i.e. F tests in AOV tables) over all treatment means, or all pair-wise comparisons (i.e. mean separation letters) include comparisons that are not of scientific interest.

- ARM notation for contrasts mimics the syntax used to select treatments for analysis. This allows ARM to determine contrast coefficients in the background and should simplify the use of contrasts for the researcher.

Appendix

- Design notes for error messaging. Kept for historical purposes.

Specification Errors

- When a comparison specification is entered in the contrast table, ARM scans the text for formatting errors.
- If no error is found, text will be displayed in normal, black text. If an error is found, the display will change to red italics, and the tooltip will contain an error message.
- In the following slides, we show different types of contrast specification errors and their associated error messages.