Exercise: Using VIT to run a randomisation test for proportions

This exercise will enable you to use VIT to run a randomisation test for an experiment where we are interested in proportions.

VIT has a Randomisation test module that is used for testing how "significant" our difference is. We can use randomisation tests to see if our experimental results could be produced just by the "luck of the draw".

The skills addressed are:

- Use VIT to re-randomise percentage data and create a re-randomisation distribution for proportions.

- Interpret the re-randomisation distribution and tail proportion.

This exercise cannot be done in VITonline yet because the randomisation tests capability has not yet been built. If you don't have VIT, reading over the exercise should still be useful, though, for some of the teaching points reinforced within it.

INSTRUCTIONS

Follow the instructions below to generate the graphs. Or you may prefer to print the instructions. If you have a problem doing the exercise, scroll down to Common questions.

Start VIT as you have done previously but this time select the **Randomisation tests** option.

- Click **Run selected VIT module**.
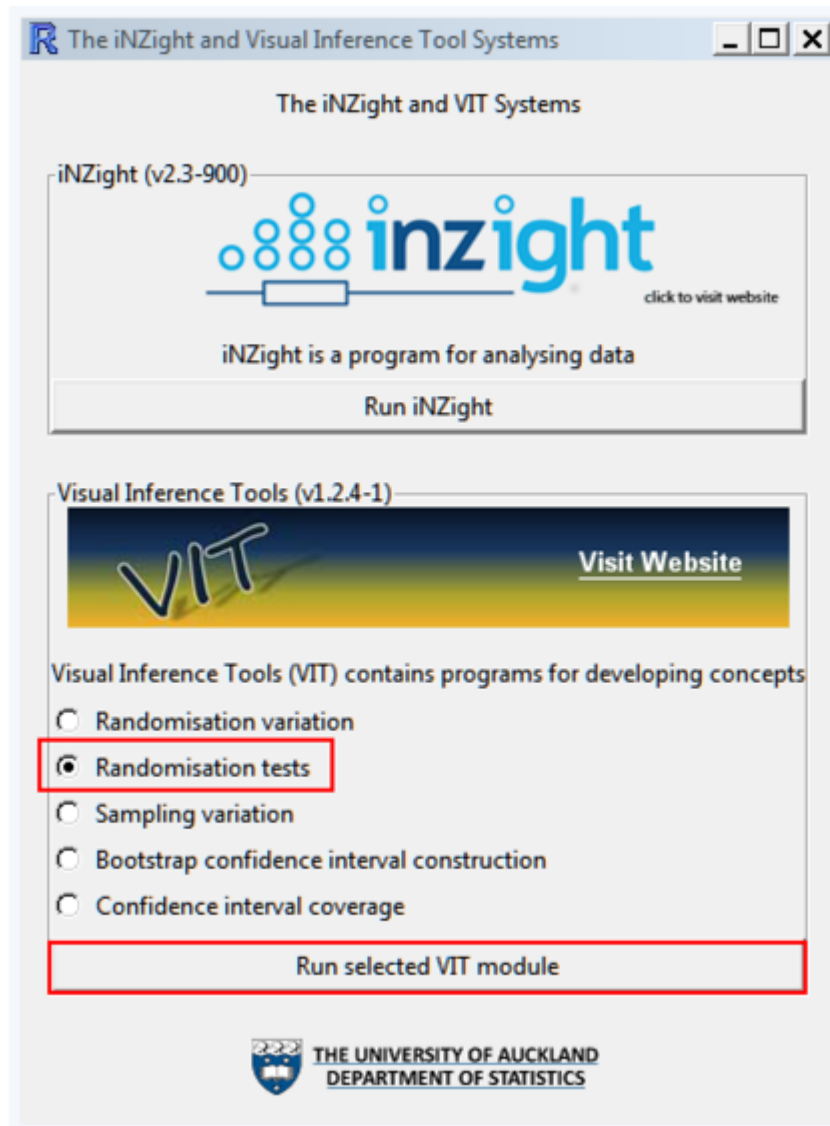
Download the **MindsetMatters.csv dataset here** or **MindsetMatters.txt dataset here** and put it on your Desktop.

Use **Import Data** and the browser window it opens up to import the **MindsetMatters** dataset into VIT from your Desktop.

This data is based on an experiment where the researcher was interested in the placebo effects of mind-set on the body. 84 female work attendants from 7

large hotels were randomly allocated to one of two groups, Informed and Control. (Those with missing values have been taken out of our dataset.)

The Informed group were told that the work they do is good exercise and meets the Surgeon General's recommendations for an active lifestyle (which was true).

The Control group were not told anything.

The participants were not told to change any behaviours, they continued working as they had done previously. Weights, BMI, blood pressures and other measures were taken at the beginning of the study and 4 weeks later.
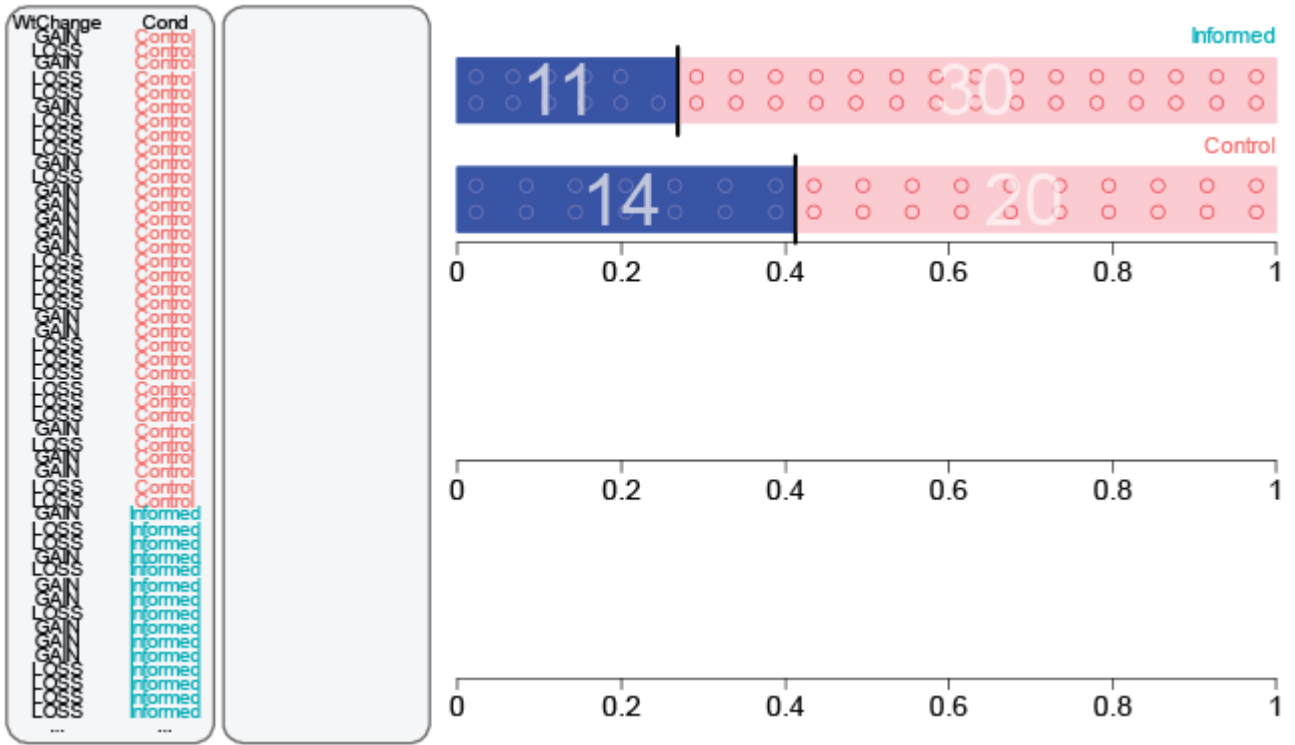
Creating a re-randomisation distribution

Drag the variable **WtChange** into the **Variable1** slot.
Drag the variable **Cond** (Informed or Control group) into **Variable2** slot.
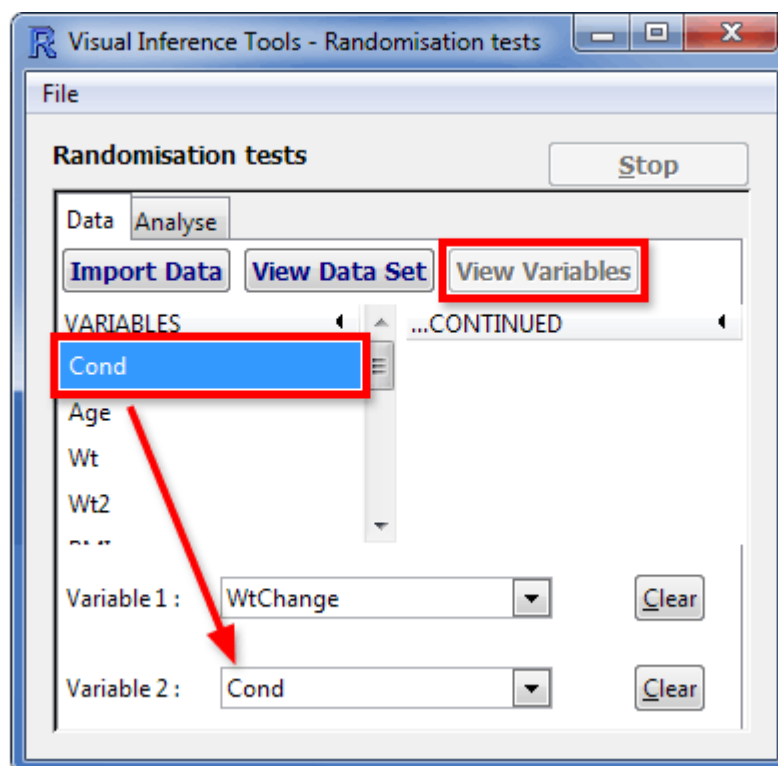
The graphics window will show 2 strip graphs with the proportion of those who gained weight in blue and the proportion who lost weight (or remained the same weight) in pink.
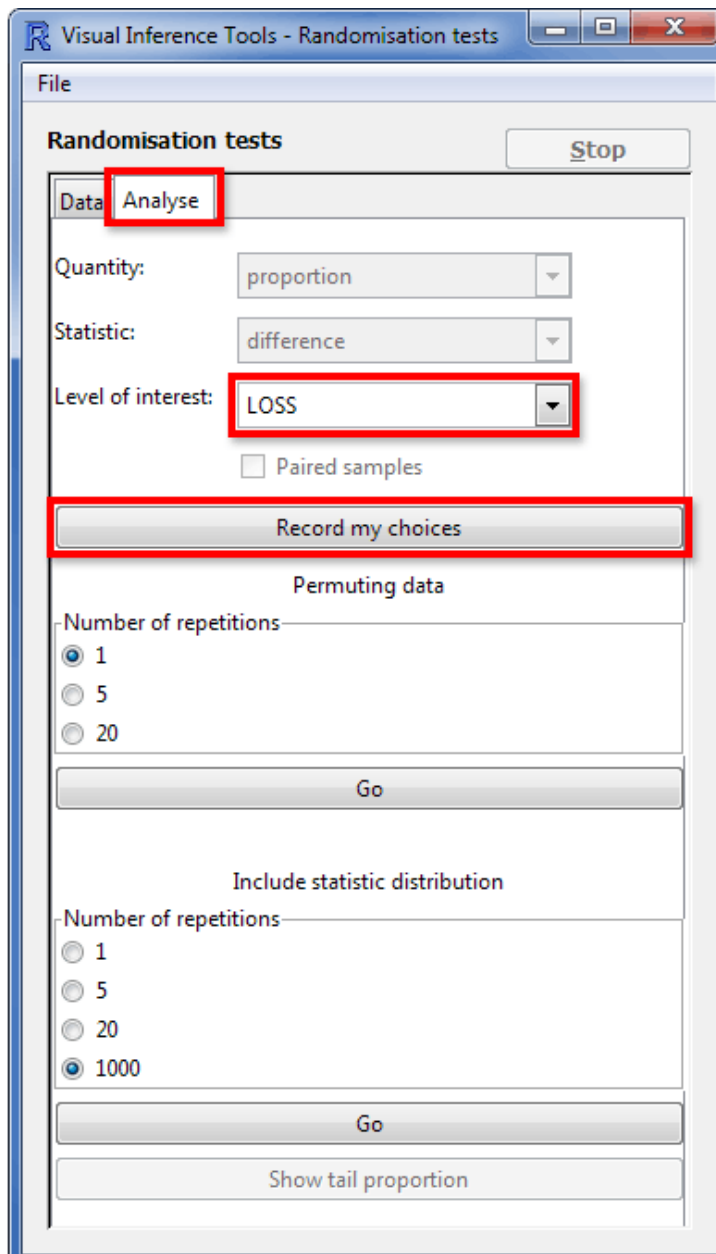
[Enlarge image](#)

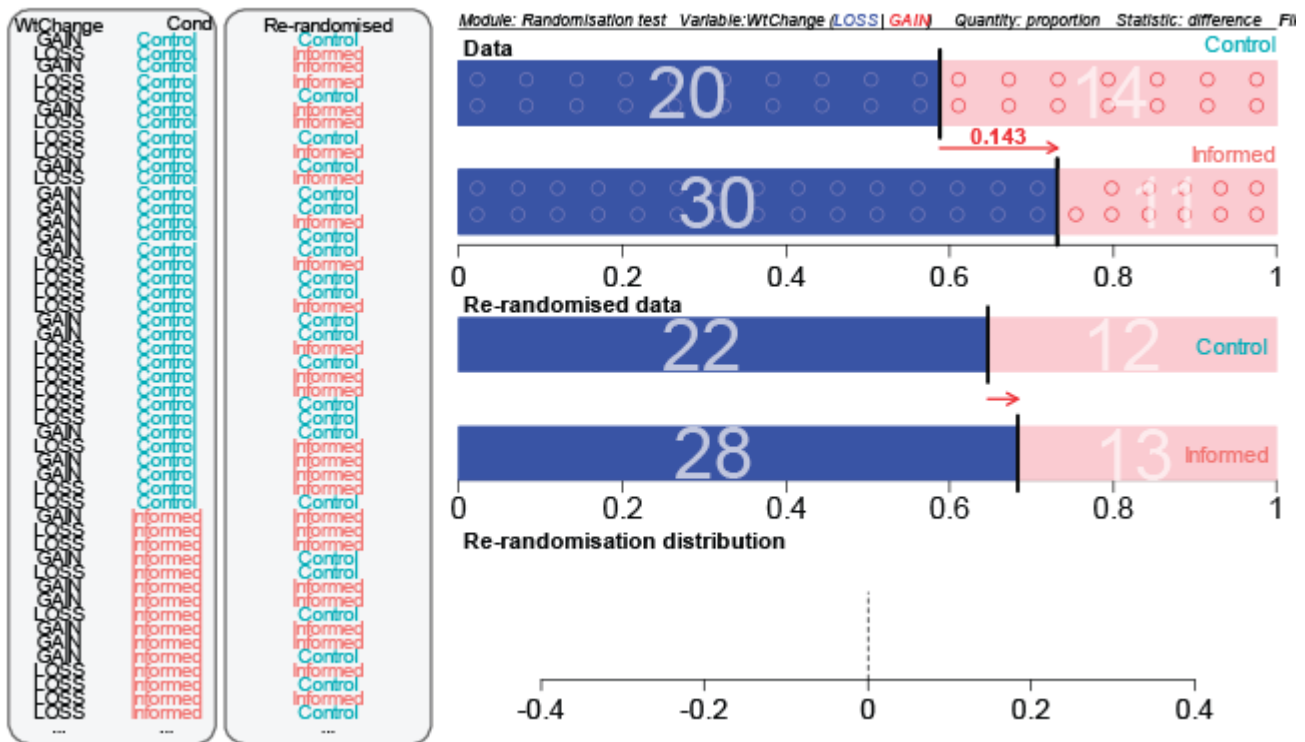*Cont.*

Now click the **Analyse** tab.

- Change **Level of interest** to **LOSS** (the other settings for Quantity and Statistic are set for proportions automatically)
- Click **Record my choices**.

Changing Level of interest to LOSS will change the graphs because they will now show the proportion of people who have **lost** weight over the 4 weeks in blue.

Now perform 1 re-randomisation of the data by mixing up which group each female attendant belonged to.

- **Select 1** repetition
- Click **Go**.

[Enlarge image](#)

## What is happening in the graphics window?

VIT is randomly relabelling the weight change for each woman "Informed" or "Control" regardless of which group they initially belonged to. This is called re-randomised data. The new proportion of female attendants who lost weight in each group is calculated and the difference of the two re-randomised proportions has been recalculated.

In the data panel of the graphics (top), what do the numbers 20 and 14 for controls, and 30 and 11 for the informed group mean? What about the new numbers in the re-randomised data window?
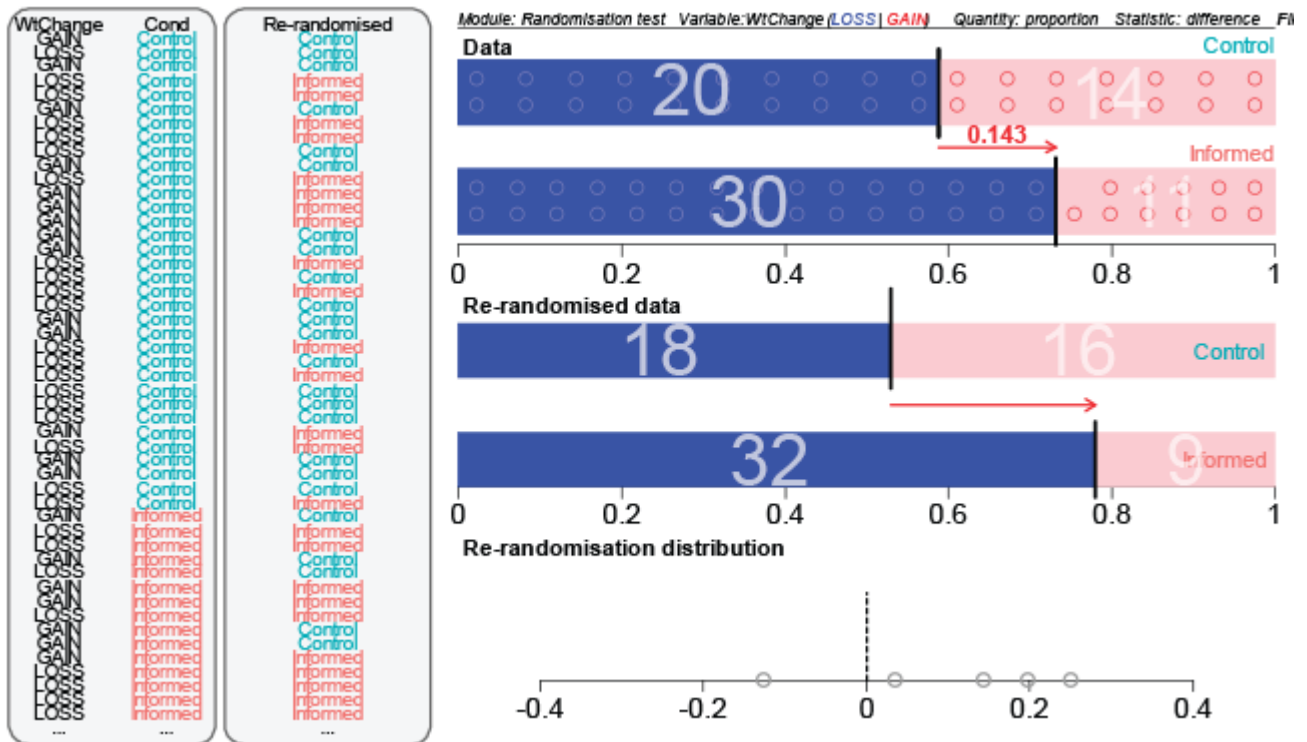
Now use the bottom part of the Analyse window that says **Include statistic distribution**.

Under Number of repetitions,

- **Select 5** repetitions and click **Go**. Take a moment to push pause and make sure you understand what is happening in this process.
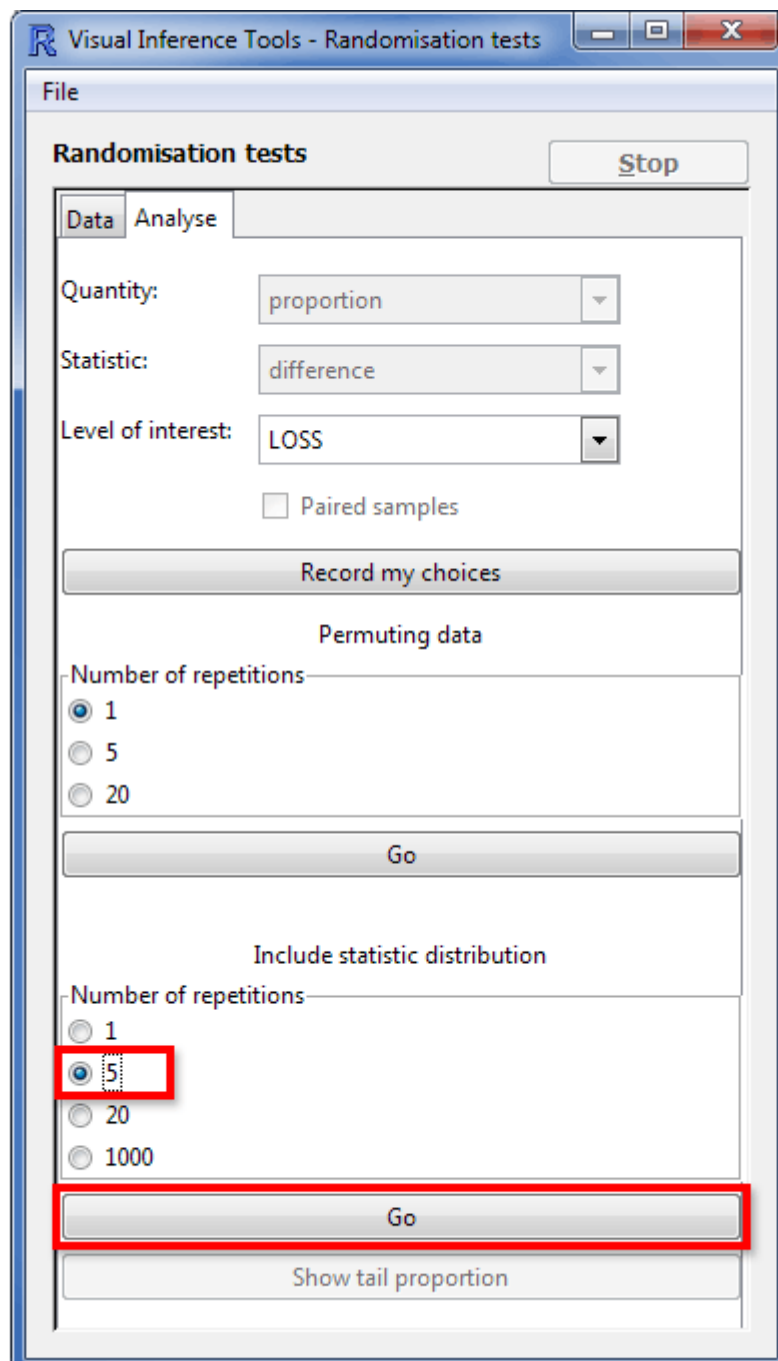
The above process of re-randomising will happen 5 times and each difference between the new randomly assigned groups Informed and Control will now be recorded in the bottom plot called the **Re-randomisation distribution**.

[Enlarge image](#)

In the bottom part of the analyse window (**Include statistic distribution**):

- **Select 1000** repetitions, click **Go**.

You will see a **Re-randomisation distribution** forming in the bottom plot. Each dot in the re-randomisation distribution is the difference between the two re-randomised group proportions (remember that the proportion is the number of female attendants in each group who lost weight over the 4 weeks).

Get the tail proportion

How many times do we get a difference of 0.143 or more between the two group proportions (Informed - control) when we reallocate female attendants to groups?

- Click **Show tail proportion**.
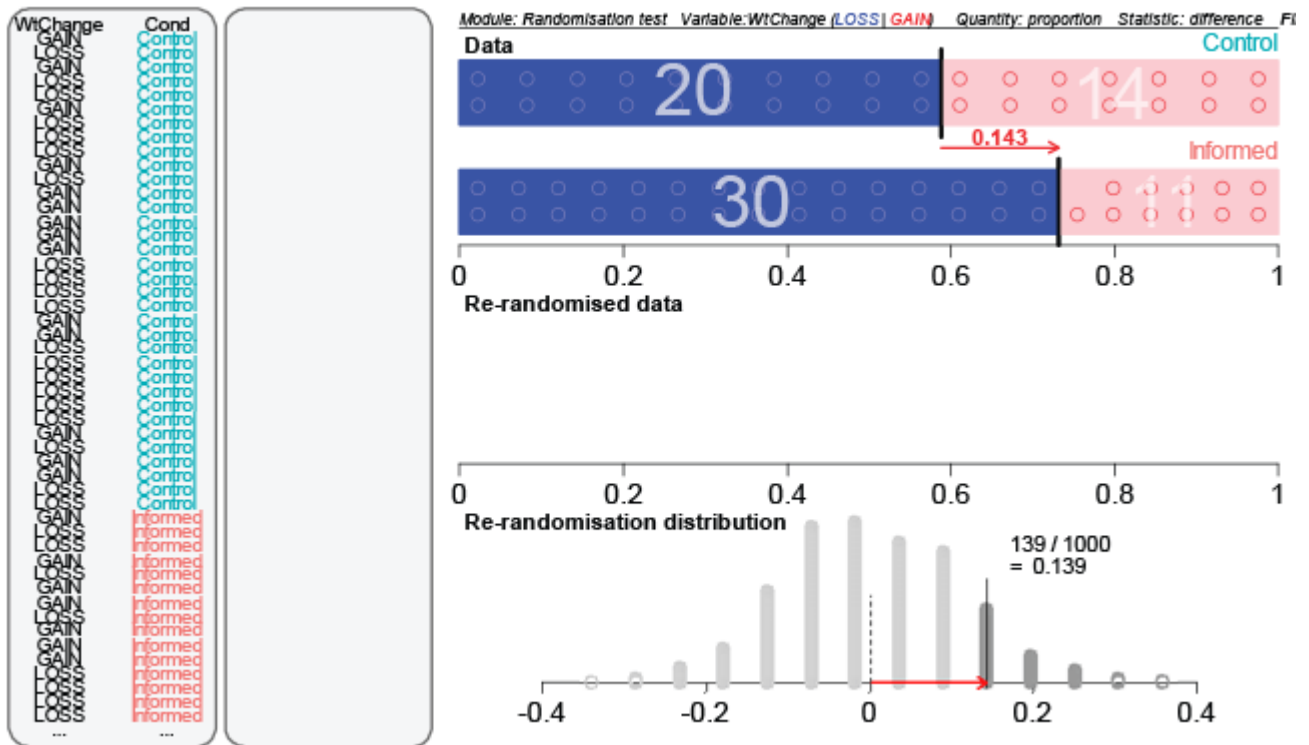
[Enlarge image](#)

What is happening in the graphics window?

The difference of 0.143 (or 14%) between the proportion of those in the Informed group and those in the Control group who lost weight is shown by the red arrow in the top plot. The difference then comes down and is drawn over the re-randomisation distribution to see how many times out of 1000 re-randomisations we got a difference the same as (or larger than) our experimental difference. In this example, the difference of 14% or larger occurred 139 times out of 1000.

Note that your values will probably be slightly different.

EXPLORE (~5 min)
Have a look at the variables **BMIChange** or **SystBPChange** in the**Variable1** slot. Run a new randomisation test and post a comment if you see anything interesting.

Common questions

Why do the data change their group allocation?

VIT is randomly reallocating the female attendants' results to the two groups to try to see how the experimental results compare to what random allocation acting alone could do. The real group allocations are stripped away from all of the **WtChange** values and the **WtChanges** are randomly (and artificially) relabelled as being in the "Informed" or "Control" group. Then VIT measures the new difference between group proportions for these new randomly formed "groups". This is one re-randomisation.

Why is my tail proportion different to the example?

The re-randomisation process is different every time VIT runs it because it randomly re-allocates the **WtChange** to groups. You should get a tail proportion that is fairly similar. If we ran it a much larger number of times (say 100,000) we would always get the same answer.

What does a small tail proportion mean?

The smaller the tail proportion, the less likely it is that the "luck of the draw" would produce a difference this big. A small tail proportion tells us that our experimental results are out in the tails of the randomisation distribution, so that random reallocation almost always gives differences that are smaller than what we got in our experiment.

What does a larger tail proportion mean?

It tells us that re-allocation often produces a difference larger than we got from our experimental results. Our experimental results give no evidence a true treatment difference exists.

When can we say that there is an effect?

This is discussed in the article that follows (Step 7.11).