DATA TO INSIGHT: AN INTRODUCTION TO DATA ANALYSIS

THE UNIVERSITY OF AUCKLAND

THE UNIVERSITY OF AUCKLAND NEW ZEALAND Te Whare Wānanga o Tāmaki Makaurau

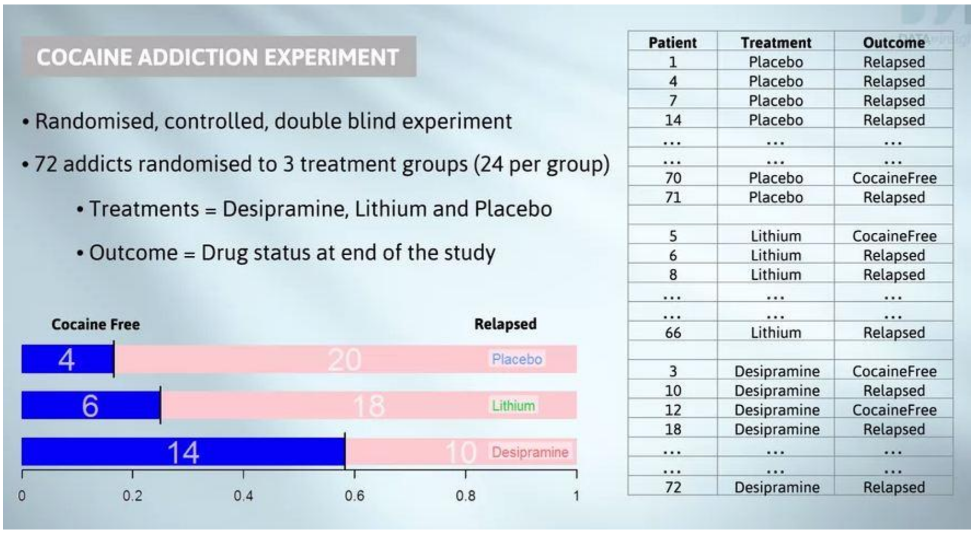
WEEK 7

PERCENTAGES AND MULTIPLE GROUPS by Chris Wild

Hello. In the last video, we introduced the randomisation test and the biggest ideas underlying its use. But we worked entirely in the context of a difference in centres. (We used means.)

And we only dealt with two groups. This time, we'll work with percentages and extend to considering more than two groups. The beauty of the randomisation test is that the underlying ideas are identical, and everything looks pretty much the same no matter what quantity you're using.

We're going to look at data from an experiment to compare the effectiveness of drug treatments for addiction to cocaine.



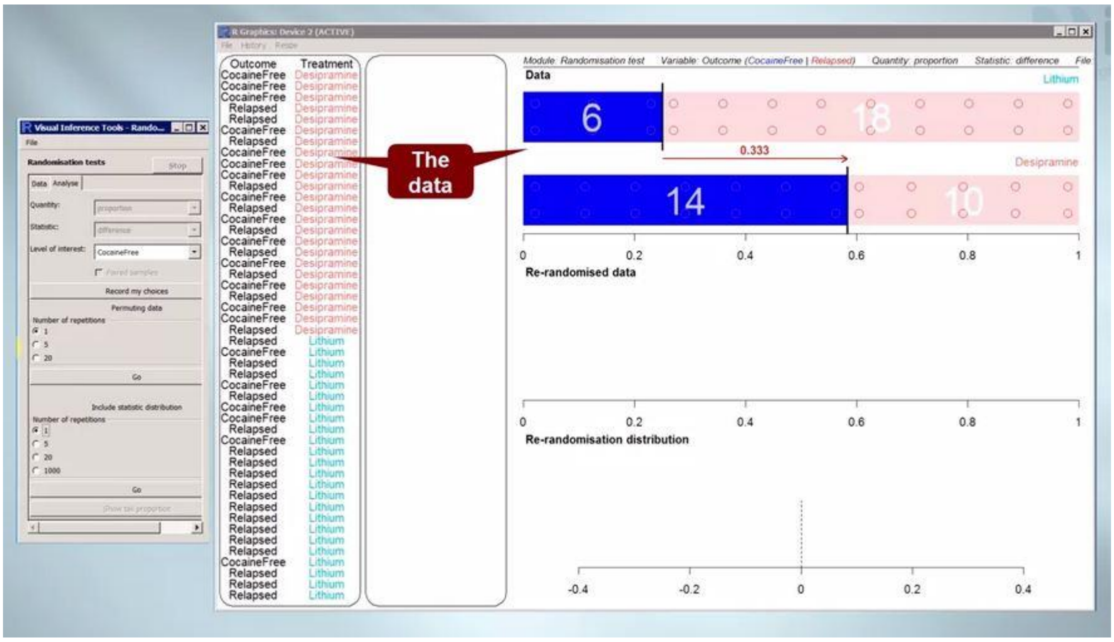
72 addicts were randomised to three treatment groups: Placebo (a dummy treatment) Lithium or Desipramine. The person-level data looks like this, which is plotted here.

We see that only 4 of the 24 people on the placebo treatment were cocaine-free. (That's about 17%.) The rest relapsed and used cocaine again. Six of those on

Lithium were cocaine-free (that's 25%) whereas nearly 60% of those on Desipramine remained cocaine-free.

On the face of it, Desipramine looks very effective compared to Lithium or Placebo. But can the differences we're seeing be explained simply by randomisation variation? We want to apply the randomisation test.

But we won't go straight to the 3-group situation. We'll see how things work for comparing two percentages first. We'll just compare Lithium and Desipramine.



Here, we're seeing the data from just the two drugs groups. On the right, they're plotted with segmented bars. The blue segments tell us how many in each group were cocaine-free, and the pink segments tell us how many relapsed.
The scale uses proportions between 0 and 1. So we see that a proportion of about 0.25 (25%) in the Lithium group remained cocaine-free, compared to nearly 0.6, (or 60%) in the Desipramine group.

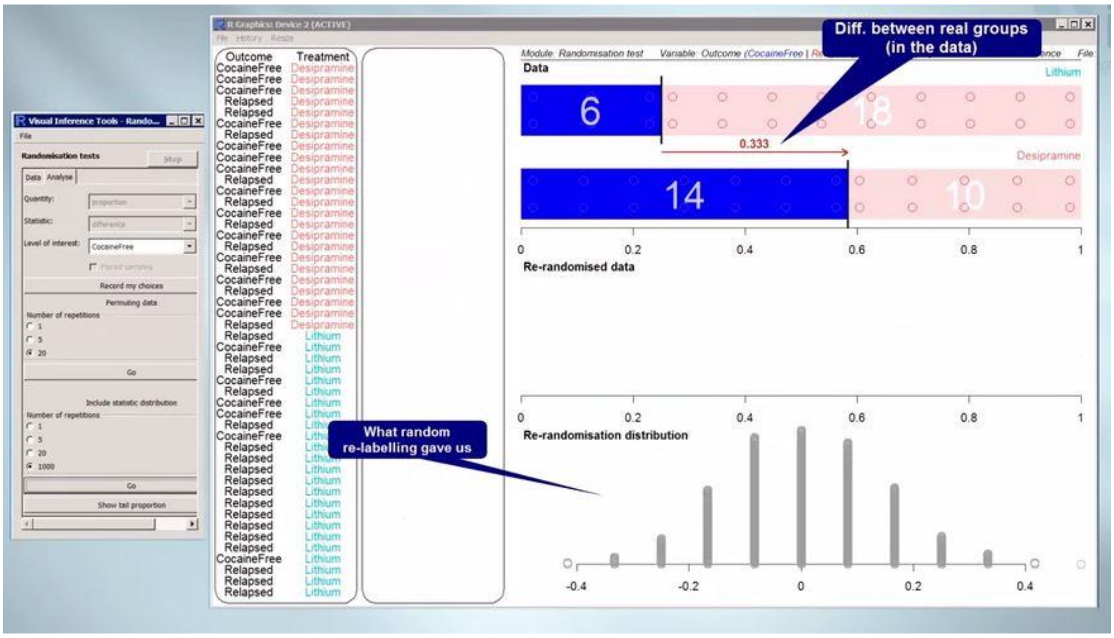
Could the luck of the randomisation draw produce a difference this big? Let's set off the animation. Here, we're using a different form of graph at the top--segmented bars instead of dotplots. But apart from that, everything we'll see is exactly as it was for means.

Here, we're copying everyone down into the middle panel, stripping off the group labels, and putting them together. Now, we're randomly relabeling them, pulling them apart into their new groups.

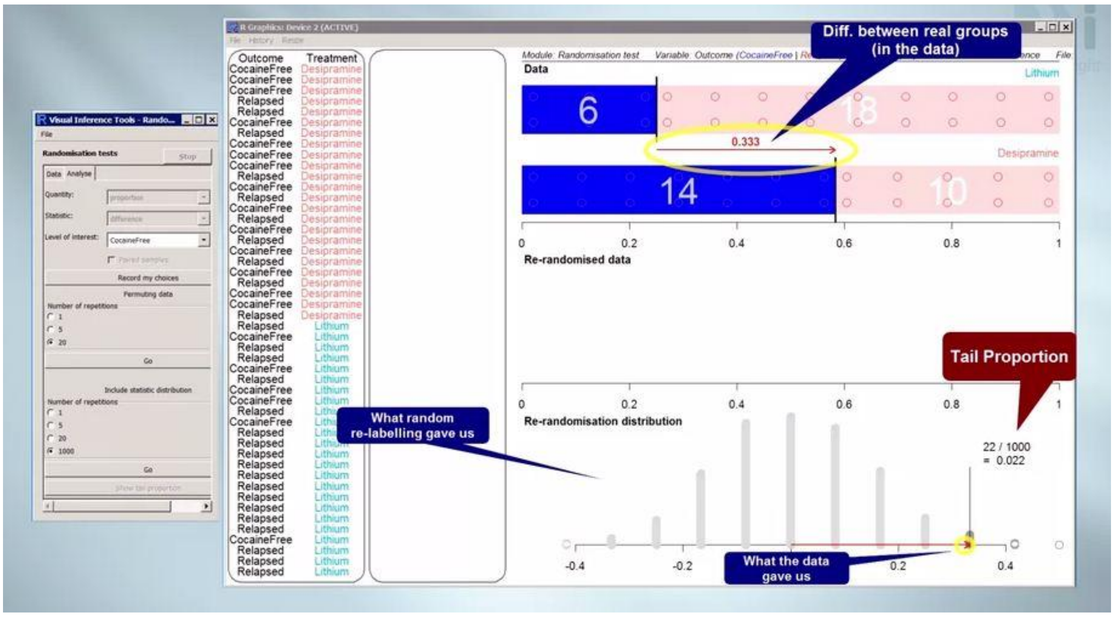
Here, we can compare the real difference (at the top) to the difference from the first re-randomization (middle). The difference from the first randomisation is actually zero. Let's do it again.

The difference in the second re-randomization is still much smaller than the one from the data. 20 times fast. Everyone I saw was smaller, but a couple were close. A reminder of how we "remember" these differences.

Do it 1,000 times.

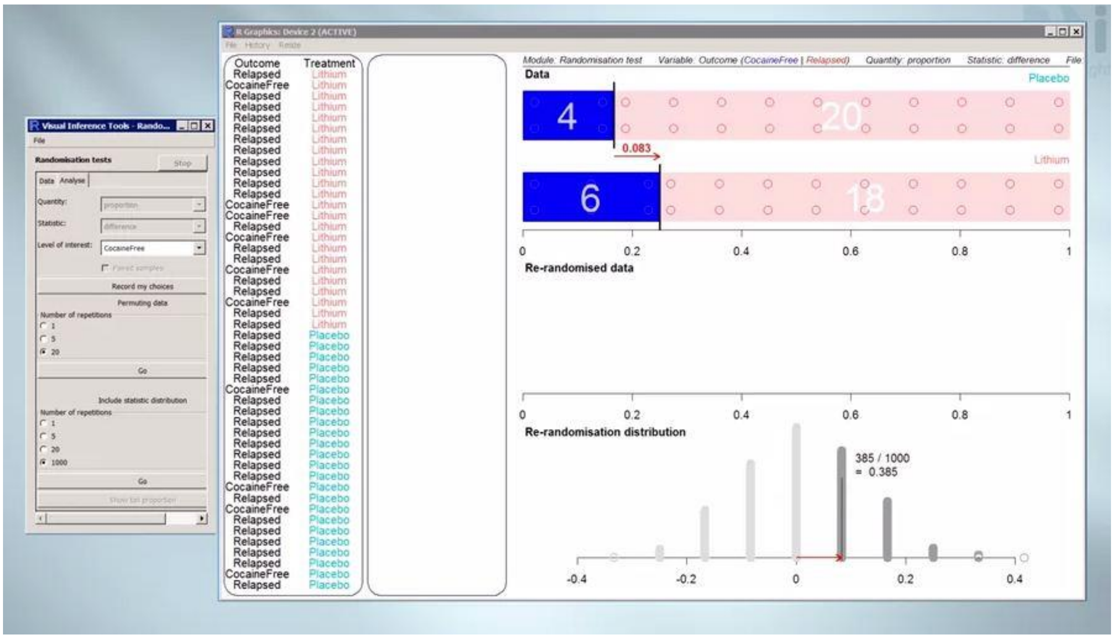


So those are all the values that random relabeling gave us. How does that compare to what the data gave us?



There it is, superimposed.Where does that fall on the randomisation distribution? It's very close to the edge. The tail proportion is 22 out of 1,000, which is about 1 in 50.

It's rare for the luck of the randomisation draw to deliver differences this big. I'm happy to conclude that I've quite good evidence that there's a real difference, and that these people were more likely to remain cocaine-free under the Desipramine treatment than they would've been under the Lithium treatment.
It doesn't always turn out like this, however.

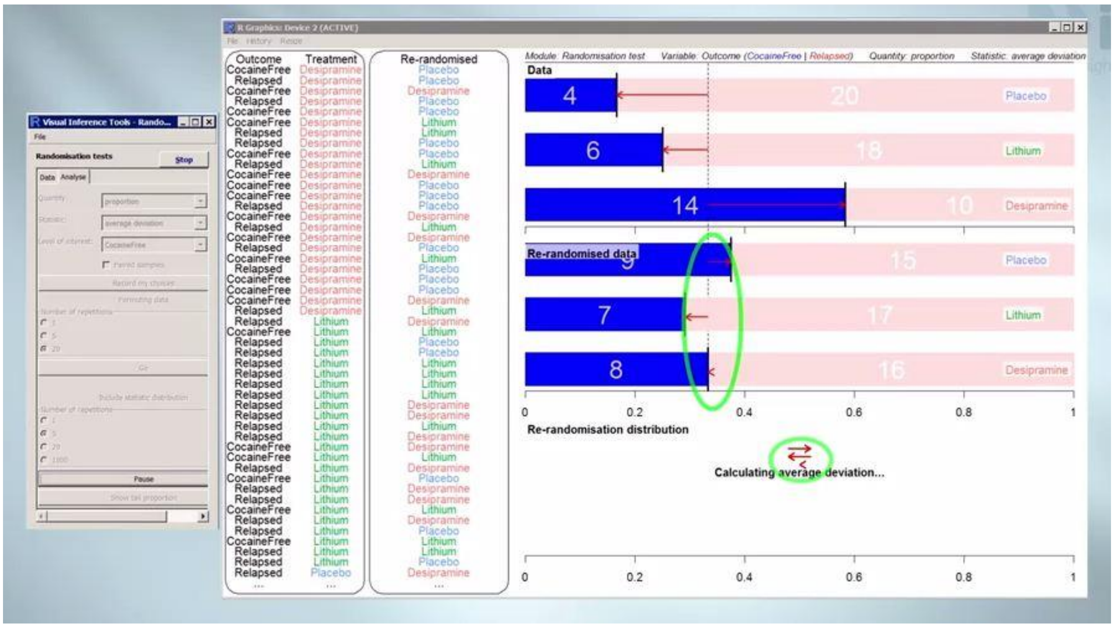


Here's the comparison between Placebo and Lithium. The luck of the randomisation draw can easily produce differences this big, or even bigger. We've got no evidence here that Lithium is a better treatment than the Placebo dummy treatment, although that shouldn't be surprising from the picture at the top.

Now, let's go back to the original problem. In a situation like this, statisticians often do a test to answer the question, "Is there anything interesting going on here at all?" Could all the differences we see be explained simply in terms of the luck of the randomisation draw?

If the answer to the latter was yes, they'd probably not bother going any further. The randomisation test still works pretty much the same in the three or more group situation. We'll highlight the differences.

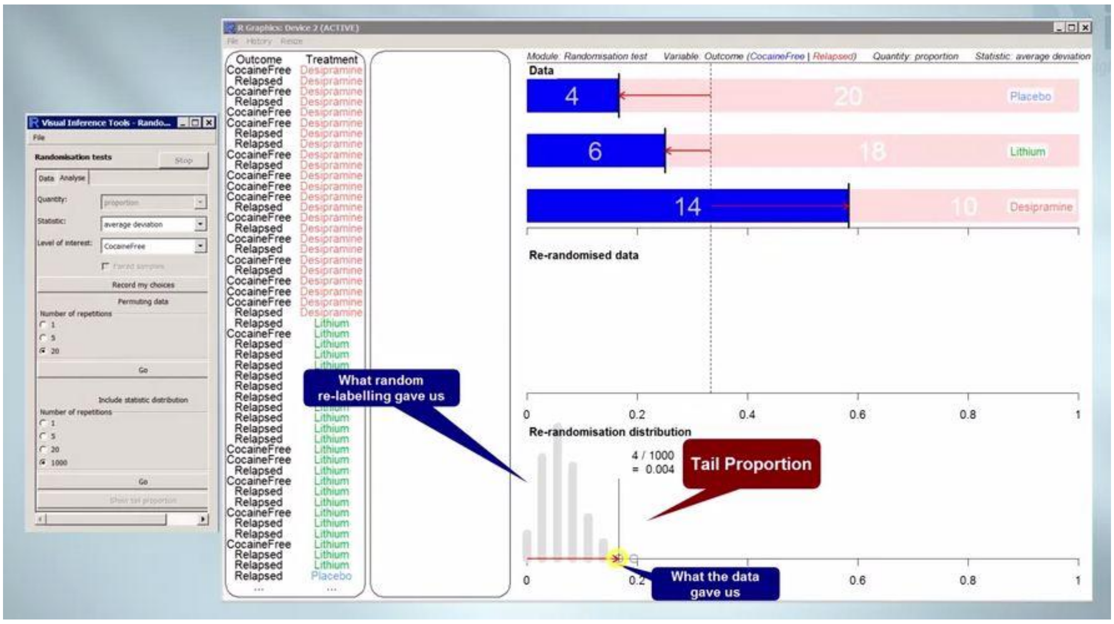
Because it's hard to draw all the possible differences on the plot, a centre line has been drawn. Its position is the proportion cocaine-free for everyone (ignoring groups). Distance arrows are drawn between the centre and each bar. The length of the set of arrows tells us about how different these proportions are. The more different they are, the longer the arrows get.



Now, we'll look at what random relabeling produces. This first set of difference arrows look similar but a bit shorter. Do it again. This set is definitely shorter. Do some faster. They were all shorter.

But how can we remember these things? We need to reduce the set of arrows to a single number measure of how long they are. Let's use the average length, which, in the software, is called the mean deviation. Here it is being calculated now.
We're going to remember this mean deviation. Do that a few times.

Now do it 1,000 times.



So here's all the deviations random relabeling gives us. How does the real one compare? Here's the arrows coming down, being averaged, dropping down. So where are we? Right near the edge.

The tail proportion is 4 out of 1000 (which is 1 in 250). The luck of the randomisation draw virtually never delivers a set of differences this big. Comparing three or more means or medians can be done in exactly the same way as we've done with proportions.

We've used every average arrow length as our combined measure of differences. Statisticians use more complicated measures of discrepancy-- chi-square statistics for proportions, and F-statistics for means. You can ask for these distance measures in VIT. They usually produce almost identical answers. That brings us to the end of this video.

