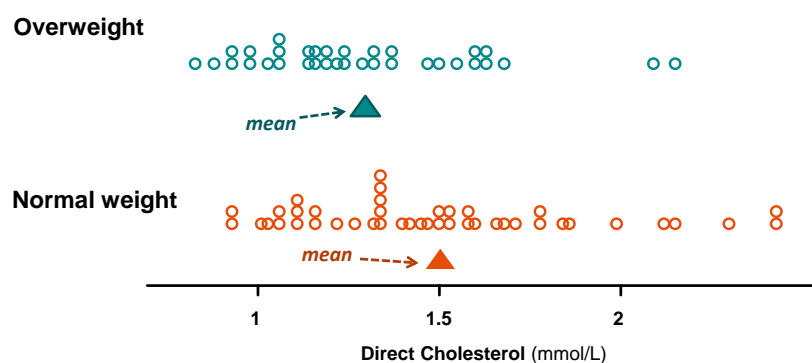


Confidence intervals for differences and visual approximations

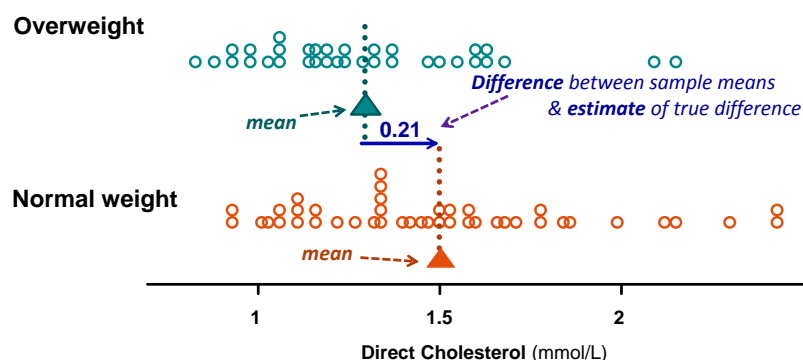
Chris Wild, University of Auckland

In this article we address misconceptions about **confidence intervals** and iNZight's **comparison intervals**. The latter merely provide visual approximations to confidence intervals for group differences that can be plotted on the original graph of the data.

Here we have data from people in their 30s from NHANES who have BMI levels classified by WHO as either **normal weight** or **overweight**. (Those classified as underweight or obese have been omitted). Is there a relationship between weight groups and direct-cholesterol levels?

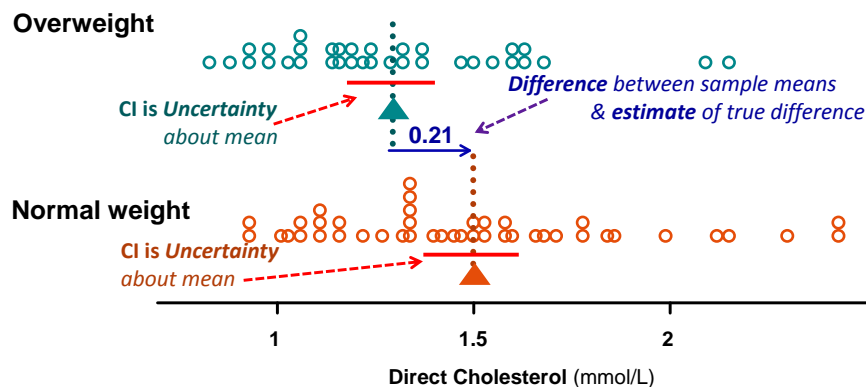


We do see a small shift towards higher values for the normal-weight group. The triangles mark the positions of the sample-means for the two groups.



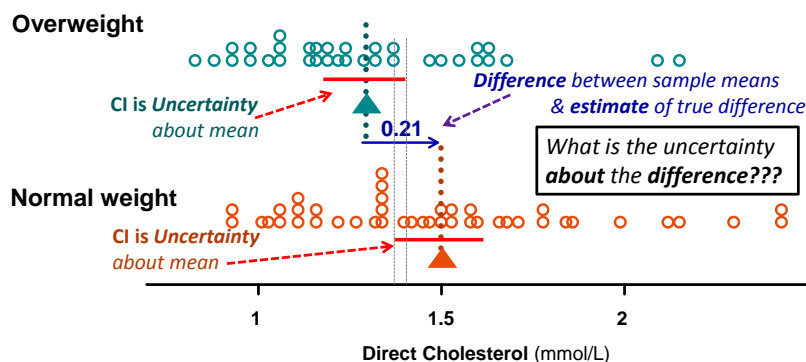
The difference in sample means gives us a point (single number) estimate of the true difference. We estimate that the true mean direct-cholesterol level for people with "normal-weight" group is larger than the mean for the "overweight" group by about 0.21 units. But what is the uncertainty around that?

We can obtain confidence intervals for the underlying true means for each group. These are easy to draw on the graph as we have done below using red horizontal lines.



Reflection: What uncertainties are being conveyed by the confidence intervals drawn on the graph above?

By putting the confidence intervals onto the graph we have represented the uncertainty in our estimates of each of the true means, but we come back to our question, **"What is the uncertainty about the difference?"** We are interested in the difference in true (population) mean levels of direct cholesterol between the two groups. To what extent does the normal-weight group have higher or lower cholesterol levels, on average, than the overweight group? The individual confidence intervals we have added to our graph in red get close to being able to answer this for us, but we will find they have problems and that's what we will now address.



The numerical values for the 95% confidence interval for the overweight group puts its true mean somewhere between 1.18 and 1.40. The interval for the normal-weight group puts its true mean somewhere between 1.38 to 1.62. These individual confidence intervals overlap and that is what we see when we look at the red lines on the graph. The overlap suggests that, in contrast to the impression given by the data, the true mean for the normal-weight group could possibly even be smaller than that for the overweight group. This would be the case if the true normal-weight mean was towards the left of the plausible range shown and the true overweight mean was close to the upper end of its plausible range.

Reflection: What does the overlap between the depicted confidence intervals suggest visually?

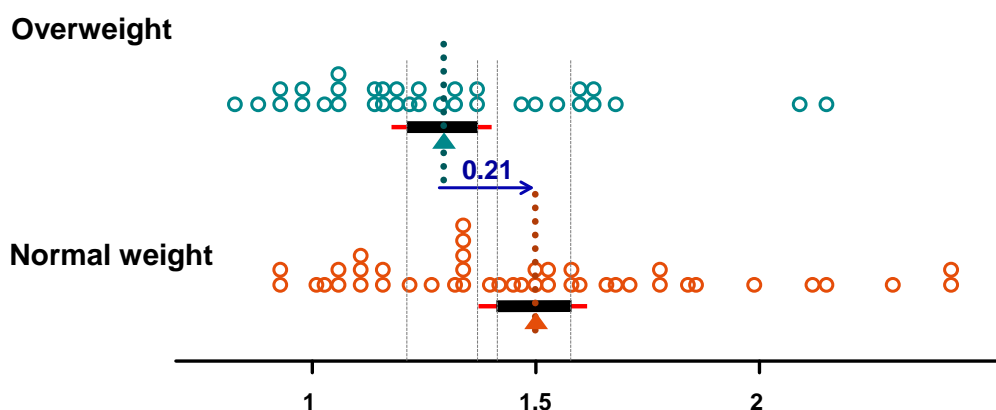
Using either bootstrapping or normal theory we can actually obtain a confidence interval for the difference in true means. (How such an interval can be constructed using the bootstrap is shown in [this movie](#)). I get a 95% confidence interval for the difference of (0.04 to 0.37). That puts the true difference somewhere between +0.04 and +0.37. The true normal-weight mean is bigger than the overweight mean by between +0.04 and +0.37.

We have two problems here. The first problem is that the visual use of the *individual uncertainties* (confidence intervals for the individual means) *overstates the level of uncertainty about the difference* (something that can be shown mathematically). The two confidence intervals for the individual means are too wide for making comparisons from.

Reflection: *What is the problem with the level of uncertainty about the difference suggested by the confidence intervals about the means drawn on the graphs above?*

The second problem is that it is very hard to convey the **confidence-interval-for-the-difference** information *on the original graph*. When we are working with several groups and want to be able to compare any two of them visually, it is impossible.

The next best thing is to **put** some sort of *an approximation on to the graph that we can see*. That is what iNZight's comparison intervals are for. They give building blocks for seeing an approximation to what we really want, the actual confidence interval for the difference.



In the picture above, the **comparison intervals** have been added in thick black. They are usually a somewhat shorter than the individual confidence intervals, thus correcting for the tendency for visual use of the uncertainties about the individual means to overstate the uncertainty about the difference. You will see that, unlike the red confidence intervals, there is no overlap between the black intervals which is consistent with the fact that the confidence interval for the difference only contains positive values (suggesting that the true mean direct cholesterol levels for normal-weight people is larger than for overweight people).

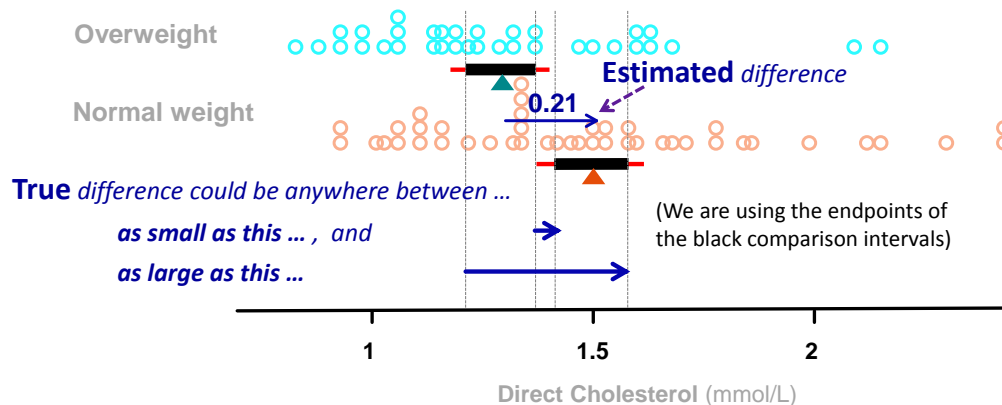
Reflection: *How do we work around the problem of the individual confidence intervals being too wide for use in making comparisons?*

The comparison intervals are only meant to be used for getting quick, initial, visual impressions about differences – so that we can see most of what is going on in the data at a

glance. We then use the output from **Get Inference** (which contains the actual confidence intervals for the differences) to confirm or correct those impressions.

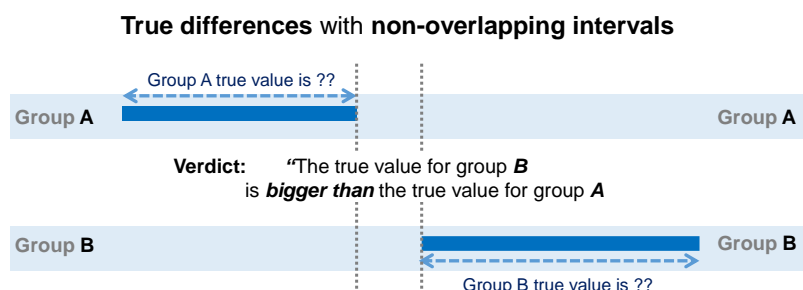
Reflection: *How are comparison intervals meant to be used? If we want more reliable information about the size of a difference, what should we look for?*

How do we read the comparison intervals visually. That was discussed in the previous movie, but the picture below shows how it works with this example ...



Visual reading of comparison intervals

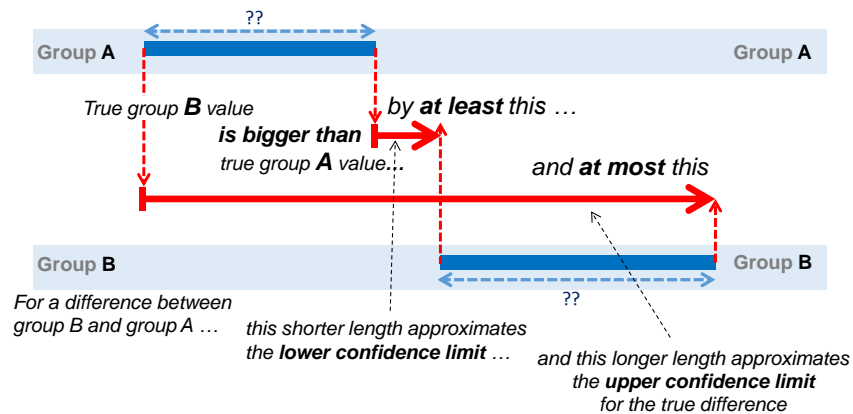
As a general pattern



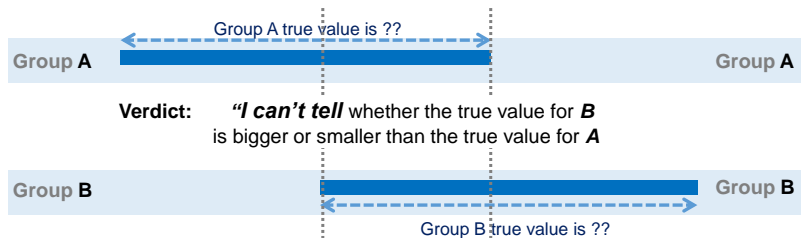
Reading off approximations to the confidence interval for the true difference ...

(The two endpoints of a confidence interval are called the lower and upper **confidence limits**)

True differences with non-overlapping intervals

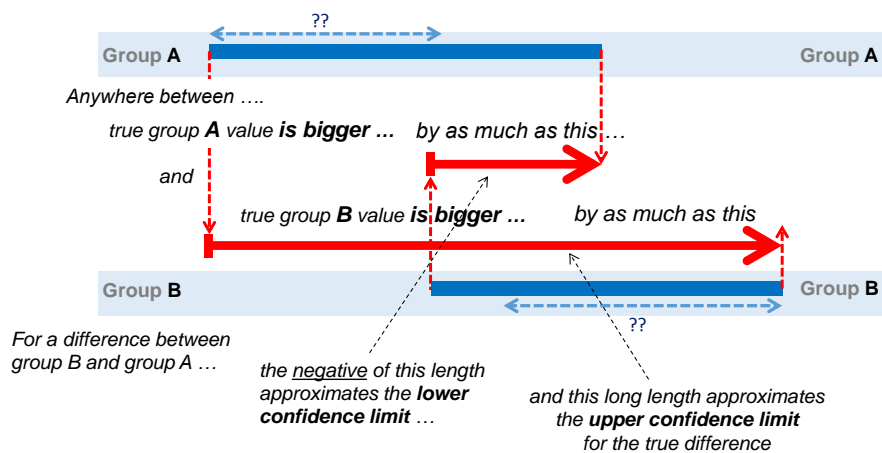


True differences with overlapping intervals



Reading off approximations to the confidence interval for the true difference ...

True differences with Overlapping intervals



Review:

(the below is written in terms of means but applies equally to medians, proportions, and other quantities)

- *What uncertainties are being conveyed by the confidence intervals drawn around means on a graph?*
- *What does the overlap between confidence intervals drawn about means suggest visually?*
- *What is the problem with the level of uncertainty about a difference in means suggested by the confidence intervals drawn about individual means?*
- *How can we work around this problem?*
- *How are comparison intervals meant to be used?*
- *If we want more reliable information about the size of a difference, what should we look for?*

© 2015 Chris Wild, The University of Auckland

(updated March 2017 by inclusion of review questions)