Appendix: hypothesis testing and confidence intervals

One way of imprecisely describing hypothesis tests is to claim that they assess whether it is plausible that the sample to hand could have been drawn from a population with parameters satisfying the null hypothesis. On the other hand, in an equally imprecise way, confidence intervals provide a range of plausible values for the parameter of interest. These aims seem so similar that it is natural to ask if there is a more formal link between hypothesis tests and confidence intervals. The answer is yes – there are quite strong links between these two entities. The link is perhaps best introduced by use of an example and the example of the thallium retention index is suitable.

The test of the null hypothesis that the mean retention index was the same for responders and non-responders gives $P=0.71$. A 95% confidence interval for the difference in mean retention index is (–37.0, 57.3). The hypothesis test yields a P value that indicates that a difference in mean retention index more extreme than that observed would occur in 71% of samples if the population mean difference were 0. In other words the data are entirely compatible with a population mean difference of 0. On the other hand, the confidence interval spans 0, again indicating that a difference in population means of 0 is compatible with the observed data. In general, when comparing means of Normal distributions, if 0 is included in a 95% confidence interval for the difference in means then the associated hypothesis test will give $P>0.05$. Conversely, if $P<0.05$, then the 95% confidence interval will not include 0.

Of course, while the data may be compatible with a population mean difference of 0, they may also be compatible with other population mean differences. Although the details have not been covered above, it is entirely feasible to test the null hypothesis that the population mean difference is x for any specified number x, not just 0. If you performed such a test and deemed that the data were compatible with a difference in population means of x because you obtained $P>0.05$, then which values of x would be compatible with your sample? The answer is that the values within the 95% confidence interval are the values of x that are compatible with your data in this sense. So, for example, if you tested the null hypothesis that the mean difference in retention index was x, for x taking any value between –37.6 and 53.7%, then the hypothesis test would yield $P>0.05$. Indeed this equivalence is true whether or not the confidence interval includes 0. The same equivalence applies between 99% confidence intervals and tests yielding $P>0.01$ and so on.

Return to Comparing two groups: I the t-test