

# TESTING TREATMENTS

## Chapter 7, 7.4 BIG ACCOUNT OF THE PLAY OF CHANCE

### WHAT DOES 'STATISTICALLY SIGNIFICANT' MEAN?

'To be honest, it's a tricky idea. It can tell us if the difference between a drug and a placebo or between the life expectancies of two groups of people, for example, could be just down to chance . . . It means that a difference as large as the one observed is unlikely to have occurred by chance alone.

Statisticians use standard levels of "unlikely". Commonly they use significant at the 5% level (sometimes written as  $p=0.05$ ). In this case a difference is said to be 'significant' because it has a less than 1 in 20 probability of occurring if all that is going on is chance.'

Spiegelhalter D, quoted in: *Making Sense of Statistics*, 2010.  
[www.senseaboutscience.org](http://www.senseaboutscience.org)

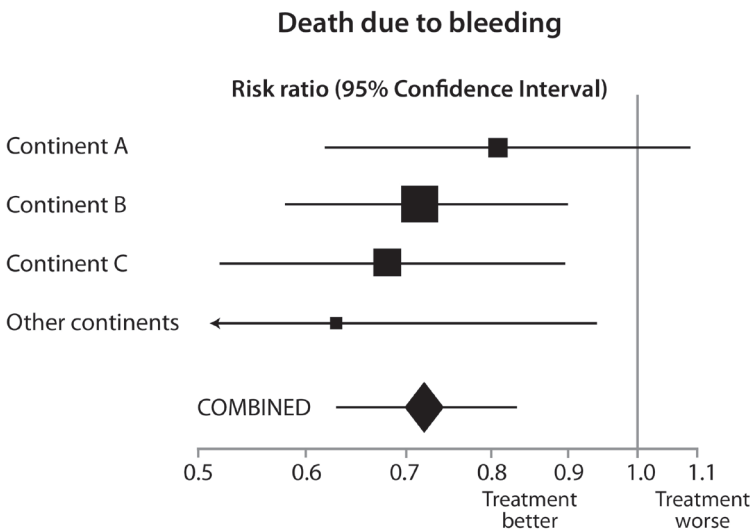
be explained by the play of chance. But that doesn't mean that it is necessarily of practical importance. If a healthy man's chance of having a heart attack is already very low, taking a drug to make it even lower may be unjustified, particularly since aspirin has side-effects, some of which – bleeding, for example – are occasionally lethal.<sup>1</sup> On the basis of the evidence from the systematic review we can estimate that, if 1,000 men took an aspirin a day for ten years, five of them would avoid a heart attack during that time, but three of them would have a major haemorrhage.

### OBTAINING LARGE ENOUGH NUMBERS IN FAIR TESTS OF TREATMENTS

Sometimes in tests of treatments it is possible to obtain large enough numbers from research done in one or two centres. However, to assess the impact of treatments on rare outcomes like death, it is usually necessary to invite patients in many centres, and often in many countries, to participate in research to obtain

reliable evidence. For example, participation by 10,000 patients in 13 countries showed that steroid drugs given to people with serious brain injuries – a treatment which had been in use for over three decades – was lethal.<sup>2</sup> In another fair test organized by the same research team, participation by 20,000 patients in 40 countries showed that an inexpensive drug called tranexamic acid reduces death from bleeding after injury.<sup>3</sup> Because these studies had been designed to reduce biases as well as uncertainties resulting from the play of chance, they are exemplary fair tests, and provide good-quality evidence of great relevance to healthcare worldwide. Indeed, in a poll organized by the *BMJ*, the second of these randomized trials was voted the most important study of 2010.

The Figure below is based on data kindly provided by the award-winning team to illustrate how, to reduce the risks of being misled by the play of chance, it is important to base estimates of treatment effects on as much information as possible. The diamond at the bottom of the Figure represents the overall result of the trial of tranexamic acid. It shows that the drug reduces death from bleeding by nearly 30% (risk ratio just above 0.7). This



**Effects of tranexamic acid on death among trauma patients with significant haemorrhage, overall and by continent of participants (unpublished data from CRASH-2; *Lancet* 2010;376:23-32).**

overall result provides the most reliable estimate of the effect of this drug, even though the estimate from centres in Continent A suggests a less striking effect (which is not statistically significant, and likely to be an underestimate of the true effect) and the estimate from centres in the ‘Other continents’ category suggests a *more* striking effect (which is likely to be an overestimate).

In rather the same way that the play of chance can be reduced by combining data from many centres in a multinational trial, the results from similar but separate studies can sometimes be combined statistically – a process known as ‘meta-analysis’ (see also Chapter 8). Although methods for meta-analysis were developed by statisticians over many years, it was not until the 1970s that they began to be applied more extensively, initially by social scientists in the USA and then by medical researchers. By the end of the 20th century, meta-analysis had become widely accepted as an important element of fair tests of treatments.

For example, five studies in five different countries were organized and funded separately to address an unanswered, 60-year-old question: in premature babies ‘What blood level of oxygen gives the greatest likelihood that babies will survive with no major disabilities?’ If the blood oxygen levels are too high, babies may be blinded; if too low, they may die or develop cerebral palsy. Because, even in these frail babies, the differences resulting from different levels of oxygen are likely to be modest, large numbers are required to detect them. So the research teams responsible for each of the five studies agreed to combine the evidence from their respective studies to provide a more reliable estimate than any one of their studies could provide individually.<sup>4</sup>

### KEY POINT

- \* Account must be taken of ‘the play of chance’ by assessing the confidence that can be placed in the quality and quantity of evidence available