

TESTING TREATMENTS

Chapter 6, 6.2.1 FAIR TESTS OF TREATMENTS

We run two risks in taking reports of the effects of treatments at face value. We could wrongly conclude that a helpful treatment is actually useless or even dangerous. Or we could wrongly conclude that a useless or even dangerous treatment is actually helpful. Fair tests of treatments are designed to obtain reliable information about the effects of treatments by (i) comparing like with like, to reduce distorting influences (biases); (ii) taking account of the play of chance; and (iii) assessing all the relevant, reliable evidence. This chapter and the next two chapters deal with these three principal features of fair tests.

COMPARING LIKE WITH LIKE

Comparisons are key

Comparisons are key to all fair tests of treatments. Clinicians and patients sometimes compare in their minds the relative merits of two treatments. For example, they may form an impression that they or others are responding differently to a treatment compared with responses to previous treatments. Sometimes the comparisons are made more formally. As early as the ninth-century, the Persian physician al-Razi compared the outcome of patients with meningitis treated with blood-letting with the outcome of those treated without it to see if blood-letting could help.

Treatments are usually tested by comparing groups of patients who have received different treatments. If treatment comparisons are to be fair, the comparisons must ensure that like will be compared with like: that the only systematic difference between the groups of patients is the treatments they have received. This insight is not new. For example, before beginning his comparison of six treatments for scurvy on board HMS *Salisbury* in 1747, James Lind (i) took care to select patients who were at a similar stage of this often lethal disease; (ii) ensured that the patients had the same basic diet; and (iii) arranged for them to be accommodated in similar conditions (see Chapter 1, p1-3). Lind recognized that factors other than the treatments themselves might influence his patients' chances of recovery. One way to make a test *unfair* would have been to give one of the treatments recommended for scurvy – say, sulphuric acid,

which was being recommended by the Royal College of Physicians of London – to patients who were less ill to begin with and in the early stages of the disease, and another treatment – say, citrus fruits, which were being recommended by some sailors – to patients who were already approaching death. This would have made sulphuric acid appear to be better, even though it was actually worse. Biases such as these can arise unless care is taken to ensure that like is being compared with like in all relevant respects.

Treatments with dramatic effects

Sometimes patients experience responses to treatments which differ so dramatically from their own past experiences, and from the natural history of their illness, that confident conclusions about treatment effects can be drawn without carefully done tests (see Chapter 5, p50-53).³ For a patient with a collapsed lung (pneumothorax), inserting a needle into the chest and letting out the trapped air causes such immediate relief that the benefits of this treatment are clear. Other examples of dramatic effects include morphine on pain, insulin in diabetic coma, and artificial hip joints on pain from arthritis. Adverse effects of treatment can be dramatic as well. Sometimes drugs provoke severe, even lethal, allergic reactions; other dramatic effects include the rare limb deformities caused by thalidomide (see Chapter 1, p4-5).

However, such dramatic effects of treatments, whether beneficial or harmful, are rare. Most treatment effects are more modest, but still worth knowing about. For example, carefully done tests are needed to identify which dosage schedules for morphine are effective and safe; or whether genetically engineered insulin has any advantages over animal insulins; or whether a newly marketed artificial hip that is 20 times more expensive than the least expensive variety is worth the extra cost in terms that patients can appreciate. In these common circumstances we all need to avoid unfair (biased) comparisons, and the mistaken conclusions that can result from them.