making the comparison unfair and the results misleading. A way to reduce differences between intended and actual treatment comparisons is to try to make the newer and older treatments being compared look, taste and smell the same.

This is what is done when a treatment with hoped-for beneficial effects is compared with a treatment with no active ingredients (a sham treatment, or placebo), which is designed to look, smell, taste and feel like the 'real' treatment. This is called 'blinding', or 'masking.' If this 'blinding' can be achieved (and there are many circumstances in which it cannot), patients in the two comparison groups will tend to differ in only one respect – whether they have been allocated to take the new treatment or the one with no active ingredients. Similarly, the health professionals caring for the patients will be less likely to be able to tell whether their patients have received the new treatment or not. If neither doctors nor patients know which treatment is being given, the trial is called 'double blind'. As a result, patients in the two comparison groups will be similarly motivated to stick to the treatments to which they have been allocated, and the clinicians looking after them will be more likely to treat all the patients in the same way.

Fair measurement of treatment outcome

Although one of the reasons for using sham treatments in treatment comparisons is to help patients and doctors to stick to the treatments allocated to them, a more widely recognized reason for such 'blinding' is to reduce biases when the outcomes of treatments are being assessed.

Blinding for this reason has an interesting history. In the 18th century, Louis XVI of France called for an investigation into Anton Mesmer's claims that 'animal magnetism' (sometimes called 'mesmerism') had beneficial effects. The king wanted to know whether the effects were due to any 'real force', or rather to 'illusions of the mind'. In a treatment test, blindfolded people were told either that they were or were not receiving animal magnetism when in fact, at times, the reverse was happening. People only reported feeling the effects of the 'treatment' when they had been told that they were receiving it.

For some outcomes of treatment – survival, for example –

biased outcome assessment is very unlikely since there is little room for doubt about whether or not someone has died. However, assessing most outcomes will entail some subjectivity, because outcomes should and often do involve patients' experiences of symptoms such as pain and anxiety. People may have individual reasons for preferring one of the treatments being compared. For example, they may be more alert to signs of possible benefit when they believe a treatment is good for them, and more ready to ascribe harmful effects to a treatment about which they are worried.

In these common circumstances, blinding is a desirable feature of fair tests. This means that the treatments being compared must appear to be the same. In a test of treatments for multiple sclerosis, for example, all the patients were examined both by a doctor who did not know whether the patients had received the new drugs or a treatment with no active ingredient (that is, the doctor was 'blinded'), and also by a doctor who knew the comparison group to which the patients had been allocated (that is, the doctor was 'unblinded'). Assessments done by the 'blinded' doctors suggested that the new treatment was not useful whereas assessments done by the 'unblinded' doctors suggested that the new treatment was beneficial.[8] This difference implies the new treatment was not effective and that knowing the treatment assignment led the 'unblinded' doctors to have 'seen what they believed' or hoped for. Overall, the greater the element of subjectivity in assessing treatment outcomes, the greater the desirability of blinding to make tests of treatments fair.

Sometimes it is even possible to blind patients as to whether or not they have received a real surgical operation. One such study was done in patients with osteoarthritis of the knee. There was no apparent advantage of a surgical approach that involved washing out the arthritic joints when this was compared with simply making an incision through the skin over the knee under anaesthesia, and 'pretending' that this had been followed by flushing out the joint space.[9]

Often it is simply impossible to blind patients and doctors to the identity of treatments being compared – for example, when comparing surgery and a drug treatment or when a drug has a characteristic side-effect. However, even for some outcomes for

which bias might creep in – say, in assigning a cause of death, or judging an X-ray – this can be avoided by arranging for these outcomes to be assessed independently by people who do not know which treatments individual patients have received.

Generating and investigating hunches about unanticipated adverse effects of treatments

Generating hunches about unanticipated effects of treatments

Unanticipated effects of treatments, whether bad or good, are often first suspected by health professionals or patients.[10] Because the treatment tests needed to get marketing licences include only a few hundred or a few thousand people treated over a few months, only relatively short-term and frequent side-effects are likely to be picked up at this stage. Rare effects and those that take some time to develop will not be discovered until the treatments have been in more widespread use, over a longer time period, and in a wider range of patients than those who participated in the pre-licensing tests.

In an increasing number of countries – including the UK, the Netherlands, Sweden, Denmark, and the USA – there are facilities for clinicians and patients to report suspected adverse drug reactions, which can then be investigated formally.[11] Although none of these reporting schemes has been especially successful in identifying important adverse reactions to drugs, there are instances where they have been. For example, when the cholesterol-lowering drug rosuvastatin was launched in the UK in 2003, reports soon began to identify a serious, rare, unanticipated adverse effect on muscles called rhabdomyolysis. In this condition, muscles break down rapidly and the breakdown products can cause serious kidney damage. Further investigation helped to show that the patients most at risk of this complication were those taking high doses of the drug.

Investigating hunches about unanticipated effects of treatments

Hunches about adverse effects often turn out to be false alarms.[10] So how should hunches about unanticipated effects of treatments be investigated to find out whether the suspected effects are real? Tests to confirm or dismiss suspected unanticipated effects