Reducing the play of chance in systematic reviews

In Chapter 7 (p91), we explained how the play of chance can be reduced by combining data from similar but separate studies – a process known as 'meta-analysis'. We used the example of five studies in five different countries organized and funded separately to address a 60-year-old quandary about what blood level of oxygen in prematurely born infants is needed to maximize the likelihood that they will survive with no major disabilities. That example illustrated how this process could be planned *before* the results of the studies were available, but the same process can be used *after* a group of similar studies have been completed.

For example, in 1974 a Swedish doctor conducted a systematic review of studies comparing the results of surgery for breast cancer with or without radiotherapy.[6] He found that, in all of the studies, women were more likely to die in the groups receiving radiotherapy. When all of this evidence was synthesized statistically using meta-analysis, it became clear that this excess mortality was unlikely to reflect the play of chance. Subsequent, more detailed analyses, based on evidence from individual patients, confirmed that the radiotherapy being used during that era did indeed increase mortality.[7] Recognizing this led to the development of safer practices.

Recognizing vested interests and spin in systematic reviews

What if the reviewers have other interests that might affect the conduct or interpretation of their review? Perhaps the reviewers have received money from the company that made the new treatment being tested. When assessing the evidence for an effect of evening primrose oil on eczema, reviewers who were associated with the manufacturer reached far more enthusiastic conclusions about the treatment than those with no such commercial interest (see Chapter 2, p18-20). However, commercial interests are not alone in leading to biased reviews. We all have prejudices that can do this – researchers, health professionals, and patients alike.

Disappointingly, people with vested interests sometimes exploit biases to make treatments look as if they are better than they really are (see also Chapter 10).[8] This happens when some researchers – usually but not always for commercial reasons –

deliberately ignore existing evidence. They design, analyze, and report research to paint their own results for a particular treatment in a favourable light. This is what happened in the 1990s when the manufacturer of the anti-depressant drug Seroxat (paroxetine) withheld important evidence suggesting that, in adolescents, the drug actually increased symptoms that prompted some of these young patients to contemplate suicide as a way of dealing with their depression.[9]

Over-reporting is a problem as well. In a phenomenon known as 'salami slicing', researchers take the results from a single trial (the salami) and slice the results into several reports without making clear that the individual reports are not independent studies. In this way, a single 'positive' trial can appear in several journals in different articles, thereby introducing a bias.[10] Here again, registering trials at inception with unique identifiers for every study will help to reduce the confusion that can result from this practice.

WHAT CAN HAPPEN IF ALL THE RELEVANT, RELIABLE EVIDENCE IS NOT ASSESSED?

Fair tests of treatments involve reviewing systematically all the relevant, reliable evidence, to see what is already known, whether from animal or other laboratory research, from the healthy volunteers on whom new treatments are sometimes tested, or from previous research involving patients. If this step is overlooked, or done badly, the consequences can be serious – patients in general, as well as participants in research, may suffer and sometimes die unnecessarily, and precious resources both for healthcare and for research will be squandered.

Avoidable harm to patients

Recommended treatments for heart attacks that had appeared in textbooks published over a period of 30 years were compared with evidence that could have been taken into account had the authors systematically reviewed the results of fair tests of treatment reported during that time.[11] This comparison showed that the textbook recommendations were often wrong because the authors