It is seven years since the idea of a monthly bulletin of bullet points of evidence Bandolier first started to take shape in November 1993. Since February 1994 there have been 81 issues of Bandolier and 10 issues of ImpAct. There have been about 1.8 million copies of Bandolier and ImpAct printed. A total of about 700,000 words have been written. People are now visiting the Internet site at a rate of over 80,000 a week, or 4 million a year, and the rate of growth is phenomenal.

**Better evidence better used**

One recent example was a researcher from a pharmaceutical company who asked “What is an NNT?” to an audience skilled in their use. Healthcare companies persist in making information about products more, rather than less, obscure. They are not learning, and that is to the detriment of the whole community. There are too many examples, like treatments for Alzheimer’s disease, where better analysis of available clinical trial information would probably make decision making easier at all levels. We could do with some cooperation.

As common are the enthusiasts and critics who latch on to some of the more obscure issues, like funnel plots, to dismiss systematic reviews and meta-analysis because it suits when budgets are tight. We show why this is wrong in this issue, as well as show how the growth in evidence-based knowledge can improve healthcare.

**N-acetylcysteine in chronic bronchitis**

One of the interesting differences between nations (and especially European nations) is that treatments considered efficacious in one are considered useless in another. One example is N-acetylcysteine for chronic bronchitis. This is widely prescribed in some parts of Europe, but not in the anglophone world. A systematic review tells us that it had benefit without harm [1].

**Search**

The strategy was wide, using three electronic databases, including the Cochrane Library, previous reviews [2] and contacting manufacturers. It was not restricted by language. Randomised comparisons between N-acetylcysteine and placebo were sought in order to answer the question of whether N-acetylcysteine had any efficacy.

**Outcomes**

Two efficacy outcomes were sought. First was the prevention of any exacerbation of chronic bronchitis. This would generally be regarded as increase in cough, sputum volume or purulence, or dyspnoea. The second efficacy outcome was patients reporting unequivocal improvement of their bronchitis symptoms. This included patients rating their treatment as good or excellent.

Information on adverse effects was also used, either as specific adverse effects, or as withdrawal from a trial because of adverse effects.

**Results**

There were eleven randomised trials with 2,500 patients randomised, and information on 2,000 patients for analysis. N-acetylcysteine doses were 400-600 mg a day in two or three oral doses. Because studies were generally of long duration (mostly 12 weeks or longer) there were dropouts, and because many of the trials were old, an intention to treat analysis was not possible. Eight trials used identical N-acetylcysteine and placebo tablets. Trials were of high quality, all scoring three or more on a scale up to five points; three scored three, seven scored four and one scored five. Nine of the eleven trials used the MRC definition of chronic bronchitis, and nearly all the patients in the studies were smokers.
Efficacy

In nine trials, no exacerbation of bronchitis occurred in 351/723 (49%) of patient taking N-acetylcysteine and 229/733 (31%) of patients taking placebo (Figure 1). The relative benefit was 1.6 (95% confidence interval 1.4 to 1.8) and the number needed to treat to prevent an exacerbation was 5.8 (4.5 to 8.1).

In five trials, 286/466 (61%) of patients rated their treatment as good or excellent with N-acetylcysteine compared with 160/462 (35%) with placebo (Figure 2). The relative benefit was 1.8 (95% confidence interval 1.5 to 2.1) and the number needed to treat for one extra patient to rate their treatment good or excellent was 3.7 (3.0 to 4.9).

Harm

In six trials, dyspepsia, heartburn or diarrhoea occurred in 68/665 (10.2%) of patients using N-acetylcysteine and 73/671 (10.9%) of those using placebo. The relative risk was 1.0 (0.7 to 1.3).

In 10 trials, withdrawal because of adverse effects occurred in 79/1207 (6.5%) of patients using N-acetylcysteine and 87/1234 (7.1%) of those using placebo. The relative risk was 0.9 (0.7 to 1.2).

Comment

This is another study from an experienced EBM group in Switzerland. The study of N-acetylcysteine complements a Cochrane review [2], but gives more useful results on treatment benefit and harm. The bottom line is that N-acetylcysteine benefits patients with chronic bronchitis without causing any treatment-related harm.

But what about the issue of clinical relevance? This again is dissected, but with a less clear cut answer. Obviously patients are doing better. Obviously patients without exacerbation consume fewer healthcare resources. There is a hint that hospital admissions are less. But maybe we lack all the tools to construct a clear-cut health economic case.

The trouble is that health economics is an inexact science, and Bandolier often feels that the less exact it tries to be the more useful it is. Health workers who deal with patients with chronic bronchitis could easily produce a simple model that would tell us whether treating patients with chronic bronchitis with N-acetylcysteine was not only good for patients, but good for healthcare systems.

Don’t forget the sting in the tail. The anglophones were wrong. The evidence for N-acetylcysteine efficacy is good. Another problem for EU harmonisation.

References:
**ALZHEIMER’S DISEASE TREATMENTS**

A major problem with studies of Alzheimer’s disease has been the diagnosis of dementia, and the choice of measurements of dementia progression. These often lack any immediate meaning, and lend themselves to dismissal as surrogate measures. Reporting these measures as means or with mean changes also tends to demean the results. A short but thoughtful analysis on a number needed to treat basis makes the case for Alzheimer treatments much stronger [1].

**Search**

The review sought any randomised, double-blind placebo comparison with more than 10 patients and of duration longer than one day in patients exclusively with Alzheimer’s disease. Outcomes where the percentage or proportion of patients who responded to treatment at some defined level were used to calculate NNTs. Crossover studies and those with enriched enrolment were excluded.

**Results**

There were five reports that met the inclusion criteria, two on rivastigmine, and one each on donepezil, tacrine and huperzine-A. The results based on ADAS-Cog movements are shown in the Table.

With higher doses the numbers needed to treat are about 5 over periods of six months or more. An improvement in the ADAS-Cog score of seven points is equivalent to reversing a typical year’s cognitive deterioration an Alzheimer’s disease.

**Table: Selected NNTs for Alzheimer's Disease treatments**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Drug</th>
<th>Dose (mg)</th>
<th>Weeks</th>
<th>Active (%)</th>
<th>Placebo (%)</th>
<th>Relative Benefit (95% CI)</th>
<th>NNT (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improved 4 or more on ADAS-Cog</td>
<td>Rivastigmine</td>
<td>1-4</td>
<td>26</td>
<td>36/242</td>
<td>38/239</td>
<td>0.9 (0.6 to 1.4)</td>
<td>N/A</td>
</tr>
<tr>
<td>Improved 4 or more on ADAS-Cog</td>
<td>Rivastigmine</td>
<td>6-12</td>
<td>26</td>
<td>58/242</td>
<td>38/239</td>
<td>1.5 (1.1 to 2.2)</td>
<td>12 (6.6 to 103)</td>
</tr>
<tr>
<td>Prevent decline of at least 7 on ADAS-Cog</td>
<td>Rivastigmine</td>
<td>6-12</td>
<td>26</td>
<td>139/149</td>
<td>138/197</td>
<td>1.3 (1.2 to 1.5)</td>
<td>4.3 (3.3 to 6.4)</td>
</tr>
<tr>
<td>Improved 7 or more on ADAS-Cog</td>
<td>Donepezil</td>
<td>5</td>
<td>24</td>
<td>23/152</td>
<td>12/153</td>
<td>1.9 (1.0 to 3.7)</td>
<td>14 (7 to 572)</td>
</tr>
<tr>
<td>Improved 7 or more on ADAS-Cog</td>
<td>Donepezil</td>
<td>10</td>
<td>24</td>
<td>38/150</td>
<td>12/153</td>
<td>3.2 (1.8 to 5.9)</td>
<td>5.7 (3.9 to 11)</td>
</tr>
<tr>
<td>Improved 4 or more on ADAS-Cog</td>
<td>Tacrine</td>
<td>160</td>
<td>30</td>
<td>26/64</td>
<td>29/116</td>
<td>1.6 (1.1 to 2.5)</td>
<td>6.4 (3.3 to 81)</td>
</tr>
</tbody>
</table>

**Comment**

This short little paper makes you think. The questions about NNTs are about what the comparison is, the duration, the dose, and most importantly how meaningful is the outcome. These authors challenge us about the meaningfulness of the outcomes. They also show that the studies have similar NNTs for other outcomes. And we should remember that ginkgo biloba produces similar NNTs for similar outcomes (Bandolier 48). The problem most people have, though, is whether the outcomes are meaningful. What is needed here is a new beginning.

First we need an independent definition about what change on what score is meaningful. That definition should include views of healthcare professionals and carers. Defining clinically meaningful harm should be part of this process.

Second we need the individual patient data from the clinical trials assessed against the new definition. That would mean that manufacturers would have to make the data available, because it is unlikely that the information is available from published sources.

Third we need a sensible assessment of costs and benefits. Not a theoretical treatise based on guesswork, but a real world assessment based on evidence.

Now that would be a first.

References:
HYPERGLYCAEMIA AND OUTCOMES AFTER MI

A randomised controlled trial showed that insulin treatment of people with MI and concomitant hyperglycaemia lowered mortality (Bandolier 48). This begs the question of what excess risk is associated with stress hyperglycaemia after a heart attack in people with and without diabetes. A systematic review [1] tells us that stress hyperglycaemia carries high additional risk for death and heart failure in diabetics and non-diabetics.

The reason or reasons why this may be so are not entirely clear, but the original observation that there was a high prevalence of glycosuria in people with a heart attack goes back 70 years.

Review

Researchers from McMaster and the Karolinska performed an extensive literature survey, but only English language reports were used. For inclusion a study had to:

- be a cohort study or clinical trial of patients admitted with acute myocardial infarction
- in which baseline glucose concentrations had been measured at or soon after admission
- and which reported outcomes of in-hospital mortality, congestive heart failure or cardiogenic shock in relation to the blood glucose levels.
- have at least 70% follow up.

Patients were regarded as diabetic if they had a reported history of diabetes. Hyperglycaemia was defined according to criteria set by the original studies.

Results

Fifteen cohort studies were eventually included in the review. Mean glucose concentrations were consistently higher in patients who died than those who did not.

Seven studies (Figure 1) reported on in-hospital mortality in patients without diabetes according to whether patients were hyperglycaemic or not (defined as blood glucose of greater than 6.0 to greater than 8.0 mmol/L). The average mortality was 25% in those with hyperglycaemia compared with 6% in those without hyperglycaemia (Table). For every five patients with hyperglycaemia after myocardial infarction one more died in hospital than would have done without hyperglycaemia.

Four studies (Figure 2) reported on in-hospital mortality in patients with diabetes according to whether patients were hyperglycaemic or not (defined as blood glucose of greater than 10 to greater than 11 mmol/L). The average mortality was 30% in those with hyperglycaemia compared with 6% in those without hyperglycaemia (Table). For every eight patients with hyperglycaemia after myocardial infarction one more died in hospital than would have done without hyperglycaemia.

Four studies (Figure 3) reported on development of congestive heart failure or cardiogenic shock in patients without diabetes according to whether patients were hyperglycaemic or not (defined as blood glucose of greater than 8.0 to greater than 11 mmol/L). The average rate was 31% in those with hyperglycaemia compared with 9% in those without hyperglycaemia (Table). For every five patients with hyperglycaemia after myocardial infarction one more developed heart failure or shock than would have done without hyperglycaemia.
rather than about causation and treatments. The Table includes a calculation of NNT as an indicator of the absolute increased risk, which was an additional 1 in 5 risk of dying or having heart failure or shock for non-diabetics, and a 1 in 8 risk of dying for diabetics. These are big risks.

Do we understand what is going on, and can we do anything about it? The paper [1] gives a good discussion of the likely cause (excess fatty acids from relative insulin deficiency), though again it could be that hyperglycaemia is simply a marker of greater myocardial damage. The cardioprotective effects of insulin and beta-blockers are also discussed.

There’s no simple answer, though. Blood glucose is clearly an important marker for morbidity and mortality. This is one occasion where the call for research to find out whether reversing the stress hyperglycaemia improves outcomes rings true. This looks increasingly important.

Reference:

One study (Figure 3) reported on development of congestive heart failure or cardiogenic shock in patients with diabetes according to whether patients were hyperglycaemic or not (defined as blood glucose of greater than 10 mmol/L). The rate was 3% in both those with hyperglycaemia and without hyperglycaemia (Table).

Figure 3: Congestive heart failure or cardiogenic shock in MI patients without diabetes (open circles) or with diabetes (filled circle)

CHF or shock with hyperglycaemia (%)

CHF or shock without hyperglycaemia (%)

Table: Summary of results in MI patients without and with diabetes

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Patients</th>
<th>with hyperglycaemia</th>
<th>without hyperglycaemia</th>
<th>Relative risk</th>
<th>NNT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>number/total (%; 95% CI)</td>
<td>number/total (%; 95% CI)</td>
<td>(95%CI)</td>
<td>(95%CI)</td>
<td></td>
</tr>
<tr>
<td>In-hospital death</td>
<td>Non-diabetics</td>
<td>136/535 25 (22 to 29)</td>
<td>84/1321 6 (5 to 8)</td>
<td>3.9 (2.9 to 5.4)</td>
<td>5.2 (4.3 to 6.6)</td>
</tr>
<tr>
<td></td>
<td>Diabetics</td>
<td>150/506 30 (26 to 34)</td>
<td>32/182 18 (12 to 23)</td>
<td>1.7 (1.2 to 2.4)</td>
<td>8.3 (5.3 to 19)</td>
</tr>
<tr>
<td>CHF or shock</td>
<td>Non-diabetics</td>
<td>55/179 31 (24 to 37)</td>
<td>531/3704 9 (8 to 10)</td>
<td>3.3 (2.7 to 4.0)</td>
<td>4.6 (3.5 to 6.6)</td>
</tr>
<tr>
<td></td>
<td>Diabetics</td>
<td>10/306 3 (1 to 5)</td>
<td>12/357 3 (1 to 5)</td>
<td>1.0 (0.4 to 2.2)</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Comment

This is an interesting adjunct to the DIGAMI trial of intensive insulin therapy in diabetics with myocardial infarction and hyperglycaemia of more than 11 mmol/L (Bandolier 48). That showed insulin therapy to reduce one-year mortality to 17% from 26% in controls. In this review of cohort studies the in-hospital death rate was 18% in diabetics without hyperglycaemia and 30% in those with hyperglycaemia.

The comparisons are obvious, but the review is about risks rather than about causation and treatments. The Table includes a calculation of NNT as an indicator of the absolute increased risk, which was an additional 1 in 5 risk of dying or having heart failure or shock for non-diabetics, and a 1 in 8 risk of dying for diabetics. These are big risks.

Do we understand what is going on, and can we do anything about it? The paper [1] gives a good discussion of the likely cause (excess fatty acids from relative insulin deficiency), though again it could be that hyperglycaemia is simply a marker of greater myocardial damage. The cardioprotective effects of insulin and beta-blockers are also discussed.

There’s no simple answer, though. Blood glucose is clearly an important marker for morbidity and mortality. This is one occasion where the call for research to find out whether reversing the stress hyperglycaemia improves outcomes rings true. This looks increasingly important.

Reference:

Implementation of secondary prevention of coronary heart disease

This conference takes place at the Kent and Canterbury Cricket Ground on January 24th. There are many reasons to attend. Most important are the topic itself, with contributions from Professor John McMurray and Dr Miles Fisher, and that the conference is part of the PRICCE programme (ImpAct 1).

The day starts with a buffet lunch at 1.00 pm and ends at 6.00 pm. It costs £50 for NHS people, and £100 for those from the private sector. Details of the programme and a registration form can be obtained from:

Edwina Watt
Hayward Medical Communications
Rosemary House
Lanwades Park, Kentford, Newmarket CB8 7PW
Tel: 01638 751515  Fax: 01638 751517
Mindstretcher on Methods

Systematic reviews and meta-analyses look for evident sources of bias or other reasons for suspecting the veracity of results of clinical trials. Funnel plots are thought to detect publication bias. Heterogeneity is thought to detect fundamental differences between studies. New evidence suggests that both of these common beliefs are badly flawed.

Bandolier generally eschews heavy stuff on complex methodological issues in systematic reviews and meta-analysis. The reasons involve their complexity, and the fact that mostly the details interest only the anoraks among us.

The trouble is that some folks seem to think that these methodological topics are cast in stone, and that any meta-analysis that had an asymmetric funnel plot or demonstrates statistical heterogeneity is so badly flawed that the treatment it reviews must be dismissed. What’s worse is that these are often reasons used to advise against prescribing treatments that are effective. So this month a light diversion into the heavy clay of maths and stats.

Funnel plots

A funnel plot is where some trial specific effect (odds ratio, relative risk) is plotted against some measure of its precision. Precision may be defined in different ways. Commonly used are the number of subjects in a trial, or some function of the standard error. If the plot is symmetric, like an inverted V, this is interpreted as demonstrated that there is probably no publication bias. If the plot is asymmetric, the interpretation is that publication bias is likely.

There is no empirical evidence to support this notion. Yet the interpretation of asymmetrical plots is often that there must be unpublished negative trials that would serve to negate the positive findings of a meta-analysis if only they could be found. Philosophers have a word for it. Falsification means that you can’t prove that its stupid though its stupidity stares you in the face. So some evidence that funnel plots are not what they seem is welcome [1].

Methods

Researchers in Hong Kong [1] examined 198 meta-analyses from the Cochrane library from 1998, excluding those with continuous variables or fewer than five trials. They then produced different funnel plots using size and using standard error. In other words, the same information from the same meta-analyses was used to produce funnel plots for each meta-analysis, but in two different ways.

Results

Of the 198 meta-analyses, 43 (22% of the total) could be construed as showing publication bias because of their asymmetry. Of these 43, 37 (86%) had a symmetrical funnel plot when a different method of plotting the results was used. As the number of trials in a meta-analysis fell, the proportion in which one of the two methods of plotting showed symmetry rose, to 100% in those with six trials. There was also a suggestion that asymmetry in at least one of the plots was present in those meta-analyses where it was least likely to be expected.

Heterogeneity

If you have a number of trials with the same sort of patients, with the same disease severity, the same intervention, and the same outcome over the same time, then what you have is a homogeneous data set. Right? Well no, actually, because heterogeneity defined statistically would call this heterogeneous 10% of the time. That’s how heterogeneity tests are set up. So if you have a homogeneous data set how well do the tests for heterogeneity hold up?

Methods

Researchers in Oxford used individual patient data to simulate clinical trials, and also to simulate 10,000 meta-analyses using different numbers of trials per meta-analysis and using different event rates. This was done for five commonly used ways of calculating heterogeneity.

Results

The most commonly used statistical tests did not produce the expected level of heterogeneity (that is, 10%) in a truly homogenous data set. They either over-estimated the level of heterogeneity (finding up to 20% of the meta-analyses heterogeneous) or underestimated it (less than 1% when the test was set to find 10%). When heterogeneity was introduced, they couldn’t detect it until the data sets were very heterogeneous indeed.

The conclusion was that homogeneity tests (what the tests really should be measuring) were of very limited use. They couldn’t detect homogeneity nor could they detect heterogeneity. The fallback is to use fixed and clearly defined inclusion criteria and fixed and clearly defined outcomes.

Comment

All types of tools have their place as we try and make the best and most appropriate sense of clinical trial data in systematic reviews and meta-analyses. But the emphasis must be on tools, not rules. Funnel plots and heterogeneity tests are not fixed in stone. They have their problems.

It is not their use that is the issue, because that may well be legitimate in any particular circumstance. The worry is that people with fixed agendas (and perhaps fixed budgets) use them as rules, and inappropriately, to try and influence prescribing decisions.

If there is any rule here, it is that clinical common sense is the best indicator of whether reviews and meta-analyses make sense. Are the patients in the trials like yours? Are the inclusion criteria sensible? Do the outcomes make sense? Are they useful? And so on. No exercise of evidence gathering was harmed by adding clinical common sense.
A question asked of Bandolier was whether all the effort put into trying to make sure that children and adults had the proper vaccinations was worthwhile. Does it have any effect on immunisation rates? The answer from a Cochrane review [1] is that it does.

**Review**

A typically thorough search strategy looked for papers with the following criteria:

- Included a patient reminder/recall system in at least one study arm.
- Reported primary research.
- Studied common national or internationally recommended childhood or adult vaccines.
- Provided immunisation coverage data.
- Were written in English.

Eligible study designs were randomised controlled trials, controlled before and after studies, or interrupted time series. Reminders were interventions that reminded patients about immunisations that were due, and recalls reminded patients about those that were overdue.

**Results**

There were 41 studies in the review. Reminders or recalls were found to be effective in 33 of them (80%). Reasons for failure to achieve benefit in the other eight were varied. They were mostly to do with design (small sample size), setting (already had very high or low immunisation rates) or because they were primarily directed at other interventions or outcomes.

The analysis examined vaccination of children and adults separately, but the main thrust of the results was much the same. Telephone reminders were the most effective (Figure), but most types of interventions were effective. Telephone reminders produced an average absolute increase of immunisation rates of 25% or more in adults and children. Other reminders produced smaller absolute increases, often of about 10%. The improvement in immunisation rates was not associated with baseline immunisation levels, which ranged from 1% to 86% in control groups at the start of the studies.

**Comment**

Reading this review reminds us how persuasive and pervasive the Cochrane Collaboration is becoming. This review will also appear on the Cochrane Library, with full details of all the studies and their results. It will therefore be a useful resource for those who want to examine how immunisation reminders are performed and managed on their patch. It might stimulate research in how to make immunisation uptake higher. It should certainly comfort those who wonder whether all the effort they put in to increase immunisation rates is having an effect. The research evidence is that it does.

For people who want to know the details of each of the individual trials, perhaps to see which applies most closely to their own situation, these will soon be available on an electronic format of the review in the Cochrane Library. The sorts of detail that will hardly ever be available in paper formats should be available with a few clicks. Your library should have it, or an organisation like doctors.net or NeLH will give you free access.

**References:**

**BOOK REVIEW**


This is a super book, and I wish I had discovered it earlier, both for myself and for my patients. It was actually recommended to me by one of those patients, after it had transformed her life in a way that neither my help, nor that of our clinical psychologist, had succeeded in doing. Such was her enthusiasm that she not only gave me the details, but was generous enough to arrange for a copy to arrive on my desk a few days later.

And who could resist even the inside front cover, which lists such intriguing gems as Distant Elephants, the case of Italian wines, Bury the Judge, and How to get rid of 90% of your worries. I turned to Distant Elephants first to read a highly relevant vignette about the risks inherent in accepting undesirable or unwanted commitments (a lecture in Edinburgh, in this case!) just because they happen to be a year ahead. The take-home message is “From far away, even elephants look small; but when you come up close, they are as large as they always are. Do not commit yourself to unimportant activities, no matter how far ahead they are”.

This is a book full of such hidden gems. I had never, for example, heard Seligman’s description of depression as “the common cold of the mind”, or Flew’s ten leaky buckets argument. Yet this is also an intensely readable, practical and accessible self-help workbook, designed to be used by ordinary people, not experts. It is, appropriately, not extensively referenced, but it has a good index, selected sources for further reading, and Chapter 2 describes the scientific background to the book’s content.

So if you (or your patients) need advice on mental keep-fit, relationships, assertiveness, anxiety, bad habits, depression, memory training, study techniques, sleep problems - well, the list goes on and on. Gillian Butler (psychologist) and Tony Hope (psychiatrist) are brilliant communicators, and present what they call a “buffet table” for our delectation. This is good psychology, well packaged and practically applied. Outstanding value at £8.99 for 638 pages of very readable stuff, and it will also enable you to find out what chunking means. Get it for yourself, and recommend it to your patients.

Tim Jack
Pain Relief Unit, Oxford

---

**BOOK REVIEW**

**Valerian for Insomnia**

Valerian is a herbal medicine that has been promoted for improving sleep. Does it work? A systematic review [1] of randomised trials provides very limited evidence for its effectiveness. It also demonstrates how clinical trials using different extracts, different outcomes, different subjects and different periods of use can produce a real mess of data.

**Review**

It sought randomised, placebo-controlled double-blind studies examining the effects of valerian preparations on sleep, without restriction of language up to May 1999. It used a range of electronic databases, including the Cochrane Library.

**Results**

Nine trials were identified. Three examined valerian over a period with consecutive use from 8 to 28 nights. Six looked at single night effects. Quality scores were generally two or less (using a score of up to five points), but were five in three trials.

One study with a quality score of five investigated 121 patients with non-organic insomnia over 28 days, with an ethanolic valerian extract of 600 mg daily. Clinical global impression was rated as good or very good by 66% of patients taking valerian compared with 26% taking placebo.

Two studies with quality scores of five investigated valerian over individual nights. One with eight volunteers with mild insomnia showed benefits from valerian over placebo. The other with 128 volunteers also showed benefits of valerian over placebo, especially in those who considered themselves habitually poor or irregular sleepers.

**Comment**

The problem is simple. There is limited evidence. Trials were often of short duration, used volunteers or patients with different criteria, and were usually methodologically poor. The positive aspect was that the three studies of high methodological quality all seemed to demonstrate some beneficial effects of valerian, though this still doesn’t add up to anything like an overwhelming case. Some adverse effects were reported.

The methodological issues are beautifully dealt with, and this is another paper that could usefully be used to train people about clinical trials and systematic reviews.

The message for consumers is that the evidence on valerian for insomnia is sparse.

References: