Testing the two-week rule

In April 2000, the UK Department of Health issued national guidelines for urgent referral by a GP of suspected cancer, where suspected was defined as either a perceived level of risk, or a hunch. This was achieved by something called a two-week rule for fast-tracking suspected referrals. The two-weeks then became a target, by which healthcare organisations could be managed.

At least that is the theory. Bandolier lacks the will to try to determine what evidence the theory was based upon, but it is interested in how well it works. Can a national policy like this make a difference, and can it make things better? Two recent studies suggest that may not be the case.

Lung cancer [1]

This was a retrospective audit of patient referrals for suspected lung cancer in Nottingham in the 12 months before and 24 months after the DoH guidelines. It abstracted route of referral, reasons, symptoms, and eventual diagnosis, as well as various markers, like time from referral to outpatient visit, diagnosis, or treatment. It involved 1,044 referrals to respiratory physicians with suspected diagnosis of lung cancer.

Results

Before the introduction of the guidelines, almost every case was an urgent referral (98%). After the introduction, referrals were either under the two-week wait referral procedures (60%), or were urgent referrals (40%). Figure 1 shows the

Figure 1: Total referrals and total lung cancers diagnosed before and after two week rule

<table>
<thead>
<tr>
<th>Year</th>
<th>Referrals</th>
<th>Cancers</th>
</tr>
</thead>
<tbody>
<tr>
<td>99-00</td>
<td>400</td>
<td>200</td>
</tr>
<tr>
<td>00-01</td>
<td>300</td>
<td>150</td>
</tr>
<tr>
<td>01-02</td>
<td>200</td>
<td>100</td>
</tr>
</tbody>
</table>
number of referrals, and the number of cancers diagnosed. The former increased substantially, whilst the number of cancers diagnosed did not change, nor did the stage at which cancers were detected.

Times between referral and diagnosis and treatment were either the same, or increased (worsened), and median times are shown in Table 1.

**Colorectal cancer [2]**

This systematic review conducted a search for published reports commenting on the effectiveness of the two-week rule in UK NHS colorectal cancer diagnostic services. Information was abstracted on the referral route, number of cancers, and cancer stage.

**Results**

Eight articles reported on about 10,000 patients referred by their GP, two retrospective and six prospective, performed between 2000 and 2003. Of these, only 24% were referred using the two-week rule procedure (range 17% to 48%), with another 24% referred as emergency cases, and 52% using other referral procedures.

Overall, 1,173 colorectal cancers were detected, about 12% of the total referred. Of patients referred under the two-week rule procedure, the proportion with colorectal cancer was 10% (range 8% to 14%).

Most patients were seen by a hospital specialist within the two-week target, and no studies reported any significant difference in the stage of colorectal cancer patients referred under the two-week rule, or by other methods.

**Comment**

Neither the lung cancer study nor the colorectal cancer review could identify any improvement in treatment using the two-week rule. In both cases almost all patients were seen within two weeks anyway, whatever the mechanism of referral. Widespread implementation of guidelines did not increase the number of cancers detected, nor the stage at which they were detected.

It would seem that for lung and colorectal cancer, we had a system that wasn’t broke, and intervention didn’t fix it. A cynic might think that one way of hitting targets is to set one up that is already being reached, and then trumpet it as a success.

As best Bandolier knows, there was little in the way of pilots, or extensive work to figure out what was needed to make an already good system better. But an awful lot of hassle was created and treasure spent for no apparent purpose.

**STATIN SAFETY: A PERSPECTIVE**

**Results**

What follows is a brief description of the results, limited to statins (simvastatin, atorvastatin, pravastatin) other than cerivastatin, and omitting discussion of fibrates.

**Rhabdomyolysis**

Cohort studies indicate a rate of 3.4 (1.6 to 6.5)/100,000 person years (1 in 29,000 per year) from cohort studies, supported by RCTs and notifications. Mortality with rhabdomyolysis is about 10%, giving a statin-specific death rate of about 1 in 300,000 per year. This is about 15 times less likely than dying in a car accident in a year.

**Myopathy**

Statin-specific myopathy from cohort studies, supported by RCTs was about 11 (4-27)/100,000 person years, a rate of 1 in 9,100 per year. It is reversible on stopping statin.
Minor muscle pain

This was no more common with statin than with placebo in RCTs, but was common, at about 5,000/100,000 person years with statin and placebo. This complaint affected about 1 in 20 of the population in the trials, irrespective of therapy.

Liver failure

The estimated risk of liver failure of someone taking statins was 0.5 per 100,000 person years, or 1 in 200,000 per year. This is about the same as people not taking statins.

Stroke

Stroke was no more common in people taking statins as not, though the rate of haemorrhagic stroke may have been higher. The advice was not to use cholesterol-lowering drugs in people who had had a haemorrhagic stroke.

Peripheral neuropathy

Cohort studies and RCTs indicated higher rates of reversible peripheral neuropathy with statins, with a statin-specific risk of about 12/100,000 person years, or a rate in 1 in 8,300 per year.

Cognitive function

No greater decline in cognitive function with statin than not.

Comment

Figure 1 tries to summarise these results in one picture, with markers of death rates from various causes in the USA in 2002 for comparison. Death from rhabdomyolysis with statin was far less common than many other causes of death, including heart disease and accidents. It would be about as likely as dying from a gunshot wound.

The figure is another use of one of the forms of the Paling Perspective Scale. Bandolier has used deaths for comparison, which is a bit unfair for myopathy and peripheral neuropathy, which are reversible on stopping statin. It would be better to have other comparators, but we are still working on that.

Because the events were by their nature rare, there was not always a large number, so to some extent the actual event rates are best guesses. For the most part, these were reasonable and conservative.

Equally interesting is why about half of people prescribed a statin stop taking it by about a year. Anecdotally, Bandolier hears that diarrhoea is responsible, but can find no studies to support this, or much at all on common adverse events with statins. Does anyone out there know better?

Reference:
**CRANBERRY TO PREVENT UTI**

Bandolier first looked at the efficacy of cranberry juice in preventing urinary tract infections in issue 6, a little while ago. Subsequent research proved to be sporadic and complicated, but a sensible clinical picture now appears to be emerging, so it seems a useful time for an update, and a quick Bandolier review.

Cranberry juice in various forms has been tested for various different effects in different situations. What we can say, with reasonable certainty, is that cranberry juice does not affect urinary tract infection rates in people with neurogenic bladder; a number of randomised trials all come to the same conclusion. Nor does cranberry juice seem to affect bacterial colonisation or infectious diseases generally in children.

The one area where there appears to be consistent benefit is in preventing urinary tract infection in susceptible women.

**Review**

Bandolier sought randomised trials of cranberry juice, or extract, in whatever form and in whatever combination with other products, in studies looking at prevention of urinary tract infection. We looked in various electronic databases, and examined bibliographies of relevant papers.

**Trials**

There have been four trials [1-4], with 779 participants, 85% of whom were women. Brief details of the trials show similarities and differences:

- **Trial 1** was conducted in elderly women (mean age 79 years) over six months, comparing 300 mL cranberry juice cocktail with a cocktail not containing cranberry juice.
- **Trial 2** was conducted in young women (mean age 42 years) who had at least two urinary tract infections in the previous year. It compared cranberry juice and cranberry extract in tablets against placebo over 12 months.
- **Trial 3** was also conducted in young women (mean age 30 years) who had at least two urinary tract infections in the previous year. It compared a cranberry juice cocktail against placebo and lactobacillus drink over six months.
- **Trial 4** was conducted in elderly patients (mean age 81 years, two thirds women) in hospital, and compared cranberry juice against placebo for less than a month.

All reported an outcome of patients with urinary tract infection, usually defined as clinical symptoms confirmed by culture (usually 100,000 colony forming units per mL).

**Results**

Combining results for urinary tract infections from all four trials, there were 42/409 (10%) in patients randomised to some form of cranberry product, and 64/370 (17%) in those randomised to a non-cranberry placebo. The relative risk was 0.5 (0.4 to 0.8), and the number needed to treat to prevent one urinary tract infection was 14 (8-45).

Individual results for the four trials are shown in Figure 1, where the dark symbol represents the shortest trial of less than one month [4]. The other studies lasted six or 12 months, and the earliest [1] also showed no difference after only one month, but consistent difference thereafter (see Bandolier 6).

Analysing for only the longer duration studies, there were 35/222 UTIs (16%) in patients randomised to some form of cranberry product, and 50/181 (28%) in those randomised to a non-cranberry placebo. The relative risk was 0.5 (0.4 to 0.8), and the number needed to treat to prevent one urinary tract infection was 8 (5-27).

**Comment**

What is interesting is the consistency of effect between these four trials, conducted in different patients, in different countries, over different periods, and using different cranberry products, although all the patients were at increased risk of urinary tract infection. In each the number of urinary tract infections was reduced by about half. In practical terms, it means that for every 10 women at increased risk of urinary tract infection who take cranberry juice daily for six or 12 months, one fewer will have a urinary tract infection.

Two further thoughts. One is about issues of validity in clinical trials, and two [1,4] have now shown less impressive effects over the short term. Any future trials should reflect the likely longer-term use, though it was entirely reasonable to conduct a short term study in hospital patients [4], in whom a long stay was unlikely.

The second is how we can express these results to women who might want to consider using cranberry juice themselves. Perhaps we should say that over six months to a year of cranberry juice use:
• 7 out of 10 women wouldn’t have a urinary tract infection anyway
• 2 out of 10 women will have a urinary tract infection anyway
• 1 out of 10 women will not have a urinary tract infection because they used cranberry juice

Maybe this expresses both the fact that it works, but that it is no miracle cure, and won’t work for everyone, always.

References:

Tiotropium for COPD

Bandolier looked at tiotropium for COPD (Bandolier 98) when the first two large trials were published. It reduced exacerbations and hospital admissions. Four years on we have a new systematic review [1], and it is interesting to see what has changed.

Systematic review

The review [1] used five search strategies, including Cochrane, and searching was up to the first quarter of 2006, making the review current. It also looked at bibliographies and requested studies from manufacturers. The review included studies in adults with stable COPD satisfying US and European criteria, where tiotropium was compared with placebo or long acting beta-agonists, that were randomised, and where the outcomes included exacerbations, hospital admission, and mortality. Only full publications were accepted.

Results

Thirteen studies with 6,078 patients were available, with 11 coming from the manufacturer (86% of patients). Most patients were men (80%), with FEV1 between 34% and 43% of predicted. Almost all studies used tiotropium at an 18 µg dose inhaled once a day. Trials were as short as one week and as long as one year, and all but one had reasonably high reporting quality, and all were described as both randomised and double blind.

Eight trials compared tiotropium with placebo (Figure 1). Overall there were fewer exacerbations with tiotropium (relative risk 0.83; 95% confidence interval 0.76 to 0.91), and the number needed to treat to prevent one exacerbation was 21 (Table 1). The NNT was similar for three large trials lasting six months or longer, and five small trials lasting six weeks or less (Table 1), but event rates were much lower in the shorter trials. However, the average number of exacerbations per week with placebo was consistent between trials at about 1.3 (Figure 2).

Tiotropium also reduced hospital admissions in the three larger studies of at least six months duration (Table 1), with a similar NNT to prevent one hospital admission, of 20 (14 to 34). In trials of six months or longer, mortality was 1.7% for both tiotropium and placebo.

Table 1: Results of trials comparing tiotropium with placebo, with outcomes of exacerbations, and of hospital admissions

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Number of Trials</th>
<th>Patients</th>
<th>Percent with outcome with Tiotropium</th>
<th>Percent with outcome with Placebo</th>
<th>NNT (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exacerbations (all trials)</td>
<td>8</td>
<td>4279</td>
<td>27</td>
<td>31</td>
<td>21 (13 to 50)</td>
</tr>
<tr>
<td>Large trials ≥6 months</td>
<td>3</td>
<td>3552</td>
<td>31</td>
<td>36</td>
<td>21 (13 to 57)</td>
</tr>
<tr>
<td>Small trials ≤6 weeks</td>
<td>5</td>
<td>727</td>
<td>4.6</td>
<td>8.6</td>
<td>25 (13 to 270)</td>
</tr>
<tr>
<td>Hospital admissions</td>
<td>3</td>
<td>3552</td>
<td>7.6</td>
<td>13</td>
<td>20 (14 to 34)</td>
</tr>
</tbody>
</table>
Most other outcomes were recorded in only a few trials, though spirometric testing in most trials consistently showed slight but significant improvement with tiotropium over placebo.

Comparison with long acting beta-agonists in two or three trials showed no significant difference for exacerbations, but lower hospital admissions with tiotropium, with an NNT to prevent one admission of 33 (17-1000), and significantly better results for spirometry.

Dry mouth was significantly more common with tiotropium than placebo or long acting beta agonist.

Comment

The estimates of efficacy from this review are broadly similar to those found in the original trials, despite there being five times more patients for the comparison with placebo. Tiotropium reduces exacerbations and hospital admissions compared with placebo, and probably against long acting beta-agonists. We could do with more comparisons with the latter to prove the point.

What was interesting was the comparison between larger, longer, trials, and smaller, shorter, trials. This could have been a source of clinical heterogeneity, or possible lack of validity, but actually the studies had a consistent rate of exacerbations with placebo, of about 1.3 per week, with only small studies varying much (Figure 2). This can provide a benchmark for knowing whether COPD candidates are like the patients in the studies.

Figure 2: The mean number of exacerbations per week of trials for placebo in individual trials. Symbol size is related to the number given placebo.

For collectors of examples of how duplicate studies can creep into systematic reviews, this paper describes how subjects in six trials excluded from this review were included at least twice in a previous review.

Reference:

percentage of responders in the comparisons in a truncated form, to emphasis the similarity between psychotherapy plus antidepressant.

Various sensitivity analyses were performed. Omitting studies that included agoraphobia made no difference, nor did comparison with different classes of antidepressant, nor did analysis by different type of psychotherapy make much difference. Additional analysis by size of trial made no difference. Sensitivity analyses like this are important, given the small size of trials generally, and the resultant scattering of results, possibly because of random chance. Figures 2 and 3 show results for individual trials comparing psychotherapy plus antidepressant with antidepressant at the end of treatment (Figure 2) and at later follow up (Figure 3).

Comment

These results appeared pretty robust. One possible quibble is that the numbers given in the paper for psychotherapy plus antidepressant compared with antidepressant alone at the end of treatment had some obvious minor problems, like numbers not adding up, and relative risks which were obviously wrong, but recalculating or omitting problematic trials made no difference. The results were robust.

Bandolier is always suspicious about results when we have only a scattering of small studies, and where there is scope for clinical heterogeneity. That was the situation here, and as we saw in Bandolier 139, that is a situation where even a systematic review is more likely to be wrong than right. That is why sensitivity analysis is important, and Figures 2 and 3 emphasise how much small trials can contribute to an overall result.

Here the results for psychotherapy plus antidepressant stood up to sensitivity analysis. The question of whether the extent of advantage over antidepressant alone is big enough, or big enough to make a cost effectiveness argument, is another matter, and a difficult call. Two years after the event some 60% of patients with panic disorder had not achieved a sustained response with any of these treatments.

Reference:

This heavy book, which was published in 2004, has only recently swum into Bandolier’s ken. We chose to review it because it has an interesting approach that some will adore, others will hate, and to which some will be indifferent. Moreover, there are electronic updates (www.evidbasedrheum.com) that make up for the inevitable delays between recency of information and publication.

It is for professionals, but has a significantly different approach in that it provides a series of summaries and decision aids for patients that address decision-making through the weight of evidence for the potential benefits and harms of interventions. It helps professionals share therapeutic decisions with patients.

Now that is incredibly difficult, because there is precious little evidence about how to do that at all, let alone well. But hey, someone has to start somewhere, and the authors deserve a big gold star just for attempting to do this across a range of rheumatological conditions. Many will find some of the decision aids and patient information fantastically useful on the back of the evidence we have. The book comes with CD, but for Windows only.

The approach is to use searches to try and answer questions (like, “Are non-pharmacological modalities effective in rheumatoid arthritis”), finding systematic reviews, randomised trials, and observational studies. In that regard, it is a more in-depth version of Clinical Evidence. It tells the reader what information there is.

Some people will quibble at aspects of the book. The levels of evidence used, for instance, gives platinum or gold gradings where there are sample sizes of 50 per group. Whatever the statistical power, random play of chance can affect both the result and the extent of any benefit or harm. Some would be more comfortable if the bars were raised. Calculating NNTs from single trials with a few tens of patients will leave some readers uneasy. They might argue that if there is inadequate or insufficient evidence, there isn’t any useful evidence.

The use of US measurement units without SI equivalents will mystify younger readers, and will be a turn off, and it is curious that the publishers couldn’t be bothered to consider the global audience.

Bandolier’s advice is to check the website. Go to a bookshop and peruse the book over a cup of coffee. Many of you will want to take it home; some others will have to stifle their screams till they get outside. It is unusual for a book to have that effect.


It is just like it says on the tin – a dictionary, which starts with the ability to pay and ends with the z-test, and provides definitions for all of them. Some of those definitions are short, and some long. But that still leaves an awful lot of terms and definitions to fill 380 pages or so.

While you might be comfortable with prevalence before you opened the book, there are few readers who would want to stand up before an audience and give a short lecture on price discrimination. Actually, price discrimination looked really interesting, and the page or so devoted to it was too short: there is much more to this than meets the eye.

The book is stuffed with terms that few of us ever come across, like extra-welfarism, or friction cost, or indifference curve, or partial equilibrium. You are thinking, so what? But if you were to read the definitions of each of these terms, with minimal neuronal responsiveness (actually, we just made that up, it’s a technical term for being awake), you would want to know more about it.

Take a moment out to think where we are, certainly in the UK, and probably with healthcare globally. The markets are taking over, and economics, and health economics in particular, is possibly the second most important influence on decisions, after politics and before or alongside effectiveness. We all should know more, and books like this are useful for the library bookshelf, to browse, and to stimulate.

Spinal manipulation reviewed

Bandolier loves systematic reviews of systematic reviews, because they can give us the quick answer about an intervention. One on spinal manipulation [1] looked for all the systematic reviews published between 2000 and mid 2005, for any indication.

Sixteen were found, on back pain, neck pain, lower back pain, headache, non-spinal pain, dysmenorrhea, infantile colic, asthma, allergy, cervicogenic dizziness, and any medical problem. Individually and collectively, there was no evidence that spinal manipulation was an effective intervention for any condition. There may be adverse events, though, and the review gives useful references for common and rare adverse events. Useful for a journal club.

Reference: