What do we think?  What do we know?  What can we prove?

Upper GI disease

If evidence-based medicine in the form of systematic reviews does anything, it makes you think. Reading any review is likely to inform, and not just in areas of ignorance: they inform also where we think we know something. Systematic reviews are also being done to answer questions that we think difficult, and using new methods to glean knowledge where previously there was only question or knowledge of the hand-me-down variety.

Bandolier has read with interest several different systematic reviews about upper gastrointestinal disease. Each examines a different aspect, and each raises perhaps as many questions as it answers.

What is the prevalence of upper gastrointestinal symptoms?

A systematic review searched MEDLINE and previous reviews on the subject [1]. Studies were included if conducted in the general population examining the prevalence of dyspeptic symptoms and reporting the period studied, sample size, response rate and definitions of questions.

Results

Ten studies were identified: six from Scandinavia, two from the USA and two from the UK. Samples were above 1000 individuals, all examined adults, and the response rate was generally above 70%. All the studies used a questionnaire, and the period over which symptoms were examined was from three months to lifetime, but most commonly one year.

Age and period made no obvious difference to the results. There were wide variations. Upper abdominal pain or discomfort affected between 8% and 54% in different studies. Heartburn and/or regurgitation affected 10% to 48% for heartburn, 9% to 45% for regurgitation, and 21% to 59% for either. Prevalence rates were lowest when the questions related only to epigastric pain (8-21%), were higher when questions included upper abdominal pain or discomfort (11-26%) and were highest when upper abdominal symptoms were included (29-54%).

Comment

Prevalence obviously depended on the questions asked, and how those questions were understood. The bottom line is that there is a lot of it about.
How effective is endoscopy in dyspepsia management?

A systematic review [2] set out to assess whether the literature could answer a number of questions relating to the effectiveness of endoscopy in improved patient outcomes or cost-effectiveness in the management of dyspepsia. Searching used three computerised databases for clinical studies (any study design) of patients with dyspepsia with information of the effects of endoscopy on:

1. Patient outcomes (symptoms, quality of life, anxiety, satisfaction)
2. Resource utilisation
3. Clinical decision making
4. Cost effectiveness

Results

For initial endoscopy in the management of dyspepsia, the weight of evidence only supports its use in clinical decision making. For the other three questions, the evidence does not support initial endoscopy.

There are important caveats, though. First, restriction of endoscopy to patients who had positive blood tests for Helicobacter pylori, who were older than 45 years, or who were taking NSAIDs would reduce the number of endoscopies by a large amount, a consistent finding in five non-randomised studies. Second, many studies had designs that were suboptimal, limiting our ability to generalise.

This review helps anyone trying to devise or review care pathways for patients with dyspepsia in primary care. The available evidence is presented, in some detail, and different weight may be placed on different aspects of the evidence depending upon local circumstances.

What happens to people with reflux oesophagitis on placebo?

What this paper sought was all trials of reflux oesophagitis with placebo controls between 1976 and 1990. It does not tell us whether or not the trials were randomised, but does tell us that only English language papers were used. To an extent, therefore, the searching and inclusion was less than satisfactory. Many more trials will have been published since 1990.

What the authors did was to extract the healing rates with placebo, both those symptom free and those with no worse than grade 1 oesophagitis. Grade 0 is given to normal oesophagus with no macroscopic damage. Grade 1 describes an oesophagus with a few areas of erythema, mucosal friability and contact bleeding. These are minor changes regarded as being present in normal oesophagus by some gastroenterologists.

Results

They found 22 studies. The healing rate for no worse than grade 1 oesophagitis varied widely between studies (Figure 1). Individual trials had healing rates between 0% and 63% at eight weeks, though only one of these trials had more than 20 patients given placebo. Overall 116/464 patients (25%) were healed at 4-6 weeks, 22/98 (22%) at 8 weeks, and 104/340 (31%) at 12 weeks or longer.

Comment

This paper may not be the best ever, but it certainly makes one think. It is not uncommon for people to make a great play of different placebo event rates between individual trials to play up or play down a particular result from a single trial. We need to be cautious about the single trial reflex (Bandolier 27), not just about extrapolating a result from one small trial, but on dismissing a result because we think a placebo response is out of line.

It also makes it possible to consider using a generalised placebo healing rate together with healing rates from systematic reviews of PPIs and histamine antagonists [4] to calculate NNTs compared with placebo. Figure 2 shows the average eight-week healing rates. Not everyone would agree with this, but given there is clinical homogeneity between the trials, the implication is that the NNT for eight week healing of erosive reflux oesophagitis with a PPI is 1.8, while that for the same outcome with an H2A is 4.4. More formal work might underpin this conclusion.

Then there’s a philosophical point. We talk glibly about placebo response rates, as if the placebo has caused the response. In this case we have no idea what would have happened without placebo. Is it credible that some psychological healing effect is at play here, or is this just the natural history of the disease?

Figure 1: Healing rates with placebo in trials of reflux oesophagitis

---

Percent of patients healed

0 100 200

0 2 4 6 8 10 12 14

Weeks of treatment with placebo
What are the consequences of long standing reflux disease?

A study of the whole population of Sweden tells us that frequent and severe symptoms of reflux over a long period are associated with very much higher risks of oesophageal cancer [5].

Study

The study found every case of cancer of the oesophagus or gastric cardia newly diagnosed between late 1994 and 1997. The Swedish system and special organisation allowed the cases to be identified rapidly and to be paired with matched controls chosen at random from the Swedish population. Patients and controls were seen by trained interviewers blinded to the intention of the study, and they were asked a number of questions about lifetime experience of heartburn and regurgitation. Symptoms occurring within five years of diagnosis were disregarded. Diagnosis of cancer was by pre-set rules, and almost all cases were reviewed by a single pathologist.

Results

There were 618 patients with cancer and 820 controls. After adjusting for a mass of different things, the results showed that oesophageal adenocarcinoma, but not adenoma of the gastric cardia or squamous cell carcinoma of the oesophagus, was highly related to symptoms of reflux. The Table shows the odds ratios for the frequency, duration and severity of symptoms. For frequency, symptom severity and for duration of symptoms there was a dose-response relationship – shown for frequency in Figure 3.

Comment

This study was very good. It shows a clear association between reflux symptoms and oesophageal cancer, and probably establishes causality. What it does not show, and is careful to express, is that treating the symptoms will prevent the cancers developing. They point out that oesophageal cancer is becoming more common, while the prevalence of symptoms is more or less unchanged and while effective treatments have been introduced. It also points out that endoscopic surveillance would swamp the system. It poses many questions, but we have to wait a little longer for definitive answers.

References:
**Breast Cancer Screening**

A number of readers have asked for *Bandolier*’s views on the recent breast cancer screening furore, in which a meta-analysis [1] asks some pertinent questions about whether screening is justified. It is difficult to jump into a complicated area and provide all the answers, but the paper itself is worth reviewing because it makes one think.

**Methods**

The paper found eight randomised trials of breast screening with mammography. Authors of the original reports supplied more information when requested. In particular key questions were asked about whether the assignment methods were concealed so that no one could foresee whether the next woman would be screened or not. Again, randomisation should produce groups with identical characteristics, like age, so that the underlying risk of breast cancer was the same in screened and unscreened groups.

**Results**

Out of the eight randomised trials examining the effects of screening on over 450,000 women, in six randomisation failed to produce similar groups. These six had other deficiencies, with four having women unaccountably missing from the analysis.

The two trials which were adequately randomised failed to show any difference for deaths from breast cancer between screened and unscreened women (Table). Screening 10,500 women for breast cancer would produce one extra breast cancer death, with confidence intervals from preventing one death in 2100 women to causing one death in 1500 women.

The other six trials were much more favourable to screening (Figure 1), and consistent in the extent of benefit. Combined they had a number needed to screen to prevent one breast cancer death of 661 (506 to 950). Pooling all eight trials produced a number needed to screen to prevent one breast cancer death of 1040 (755 to 1672).

**Comment**

The review makes an extremely good case that six of the breast cancer screening trials were flawed. In some cases, they seem so flawed as to be unreliable. This is particularly important when the absolute benefit of screening is so small. The percentage of women dying of breast cancer was always below 1% in the studies, and the difference in breast cancer death rates between studies (0.23 to 0.78%, average 0.43%) was many times greater than the difference effected by screening (-0.02 to 0.23%, average 0.1%). Figure 2 shows this graphically. It also suggests - it can be put no stronger than this – that reduced breast cancer deaths because of screening are only apparent when the baseline death rate is perhaps 0.4% or higher.

**Table: Breast cancer deaths in screened and unscreened women**

<table>
<thead>
<tr>
<th>Trials</th>
<th>Unscreened (%)</th>
<th>Screened (%)</th>
<th>Relative risk (95%CI)</th>
<th>Number needed to screen (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adequately randomised</td>
<td>177/66150 (0.27)</td>
<td>183/66013 (0.28)</td>
<td>1.04 (0.85 to 1.25)</td>
<td>-10570 (2138 to -1522)</td>
</tr>
<tr>
<td>Inadequately randomised</td>
<td>725/142052 (0.51)</td>
<td>654/182179 (0.36)</td>
<td>0.75 (0.67 to 0.83)</td>
<td>661 (506 to 950)</td>
</tr>
<tr>
<td>Pooling all trials</td>
<td>902/208157 (0.43)</td>
<td>837/248192 (0.34)</td>
<td>0.80 (0.73 to 0.88)</td>
<td>1040 (755 to 1672)</td>
</tr>
</tbody>
</table>
Giving the correct answer on breast screening is difficult. The evidence on overall reduction in death rate from randomised trials and from epidemiological studies falls short of showing much evidence of effect. The trials are suspect. The enthusiasts insist that the goalposts have moved, and that we may be playing a different game. The morbidity caused by numerous false positives remains relatively unstudied.

Perhaps the best we can do for now is to tell women the truth and let them decide. Tell them that one woman in every 1000 who undergoes breast screening may be prevented from dying from breast cancer, but there may be no benefit at all. Women should demand the best unbiased evidence on the benefits and harms of screening.

Reference:

ARE PNEUMOCOCCAL VACCINES EFFECTIVE?

A newly published systematic review [1] concludes that vaccination with pneumococcal vaccines reduces the risk of infection by very significant results and is effective in elderly people. Are these results credible?

There was a considerable effort put into searching, including searching Index Medicus back to 1938. Thirteen studies were identified up to November 1996.

Problems

The first problem is completeness. Since 1996 six more studies have been published, mostly in elderly people, or people with chronic disease in industrialised countries. One examined HIV infected individuals in Uganda.

The second problem is randomisation. Two of the early studies were only quasi-randomised (alternate allocation). One of the newly published studies is also quasi random (allocation by year of birth).

The third problem is patient populations. Three of the early studies (on South African gold miners and New Guinea highlanders) are hardly representative of people who may be treated in industrialised countries. Miner in particular were younger, and lived in tightly-packed conditions where rates of pneumococcal disease may be higher.

The fourth problem is size. Many of the studies were small, especially as some outcomes, like bacteraemia, were vanishingly small. The total number of events may be subject to the random play of chance.

The fifth problem is outcomes. A variety of outcomes may be reported – from bacteraemia, all-cause pneumonia, pneumococcal pneumonia, pneumonia death or lower respiratory tract infection.

Question

What should the question be, then, when asking whether pneumococcal vaccination is effective? Obviously we want evidence on populations similar to those we may wish to vaccinate in our community – which means elderly, or institutionalised people, or those with chronic diseases. We want the least biased evidence from randomised trials. We also want to know those outcomes most important to us – and in this case perhaps all the outcomes identified above.

Analysis

Bandolier has therefore looked at nine relevant trials. Two from the 1940s and one published in 1999 have not been included because they were not properly randomised. The studies in the 1940s had a positive conclusion. The study reported in 1999 from Finland in about 25,000 elderly people was negative.
Results

The results are shown in the Table.

All-cause pneumonia

There were 480 people with all-cause pneumonia in 6,514 people in four studies (Figure). Vaccination reduced the incidence non-significantly from 7.56% to 7.18%. The number of people needed to be treated to prevent one all-cause pneumonia was 260, but the confidence interval included vaccination causing more pneumonia.

Figure: Randomised trials of pneumococcal vaccines with all-cause pneumonia as outcome

Pneumococcal pneumonia

There were 368 people with pneumococcal pneumonia in 21,156 people in six studies. Vaccination reduced the incidence from 1.89% to 1.59%. The number of people needed to be treated to prevent one case of pneumococcal pneumonia was 325, but the confidence interval included vaccination causing more pneumococcal pneumonia.

Table: Main outcomes of randomised trials of pneumococcal vaccines in industrial countries

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Number of trials</th>
<th>Number of patients</th>
<th>Percent affected without vaccine</th>
<th>Percent affected with vaccine</th>
<th>Relative risk (95%CI)</th>
<th>NNT (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All pneumonias</td>
<td>4</td>
<td>6,514</td>
<td>7.56</td>
<td>7.18</td>
<td>1.01 (0.85 to 1.19)</td>
<td>260 (60 to 113)</td>
</tr>
<tr>
<td>Pneumococcal pneumonias</td>
<td>6</td>
<td>21,156</td>
<td>1.89</td>
<td>1.59</td>
<td>0.85 (0.69 to 1.04)</td>
<td>325 (152 to -226)</td>
</tr>
<tr>
<td>Pneumonia-related death</td>
<td>8</td>
<td>22,559</td>
<td>1.10</td>
<td>1.00</td>
<td>0.93 (0.72 to 1.20)</td>
<td>910 (266 to -645)</td>
</tr>
<tr>
<td>Pneumococcal bacteraemia</td>
<td>3</td>
<td>927</td>
<td>1.28</td>
<td>0.66</td>
<td>0.49 (0.12 to 1.96)</td>
<td>161 (53 to -157)</td>
</tr>
</tbody>
</table>

Pneumococcal death

There were 230 people who died because of pneumococcal pneumonia in 22,559 people in eight studies. Vaccination reduced the incidence non-significantly from 1.1% to 1.0%. The number of people needed to be treated to prevent one death from pneumococcal pneumonia was 910, but the confidence interval included vaccination causing more deaths.

Pneumococcal bacteraemia

There were nine people with pneumococcal bacteraemia in 927 people in three studies. Vaccination reduced the incidence non-significantly from 1.28% to 0.66%. The number of people needed to be treated to prevent one case of pneumococcal bacteraemia was 161, but the confidence interval included vaccination causing more pneumococcal bacteraemia.

Comment

There are two main conclusions to be drawn from this. The first is that the weight of evidence is that polysaccharide pneumococcal vaccines have yet to be shown to work in the types of people given them in industrialised countries. The only real evidence that they do comes from two improperly randomised studies from the 1940s. A similarly improperly randomised but large study done in the 1990s in Finland showed no benefits. Much resource is put into increasing pneumococcal vaccination in at-risk groups. Perhaps it is time to consider whether this is doing more harm than good.

The other point is one of consistency in systematic reviews and meta-analysis. The review of breast cancer screening made great play of improper randomisation affecting results. This review of pneumococcal vaccination [1] depended largely on data from improperly randomised trials for its positive conclusions. Part of the ethos of evidence-based medicine is not choosing trials to suit your own biases, but this can be hard.

There are perhaps two other useful insights here. This study cut-off its search in 1996, but was published more than three years later, by which time six more studies had been published which may have been included. Out-of-date information may be dangerous, as here, when all six recent stud-
ies showed pneumococcal vaccines to be ineffective. The second insight is to beware of studies that report treatment effects in relative terms – an X% reduction in risk. To be properly informed we need to know the absolute reduction. Only then will we have any idea how much effort we have to expend to get a particular result.

Reference

ACUPUNCTURE TO STOP SMOKING

This is by way of a late millennial reminder for those still smoking cigarettes. Some may be tempted to try acupuncture to help stop. But does it work? A first-class systematic review says not [1, 2]. Actually there are two reviews, one [1] recently published in the Cochrane Library and dates June 1999. The other [2] is also recently published, but its completion date is in 1997. The paper took two years from acceptance to publication, an interesting comment on the information age! We use information from the more contemporary review.

Search

Typically broad for this group, it included nine databases. Trials had to be randomised comparisons of acupuncture with sham acupuncture, another intervention, or no treatment. Various techniques were examined and information obtained on that and other critical issues to acupuncturists. The only outcome they were interested was complete abstinence from smoking early after treatment and at six and 12 months.

Results

There were 20 controlled studies for analysis. At early times, there was no difference between acupuncture and sham acupuncture (Figure 1, Table 1) with 2069 patients. There was no difference in abstinence rates at six months with 719 patients (Figure 2) or 12 months (Table 1). Where there was useful data on comparison with no treatment, there was no difference in abstinence rates at six months.

There was few trials comparing acupuncture with no treatment controls, and these gave inconsistent results. There was no difference between acupuncture at the ear and at other sites.

Table: Acupuncture for smoking cessation compared with sham acupuncture

<table>
<thead>
<tr>
<th>Time</th>
<th>Number of trials</th>
<th>Number of patients</th>
<th>Relative benefit (95%CI)</th>
<th>NNT (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early</td>
<td>14</td>
<td>2069</td>
<td>1.1 (0.99 to 1.3)</td>
<td>27 (13 to -710)</td>
</tr>
<tr>
<td>Six months</td>
<td>7</td>
<td>719</td>
<td>1.26 (0.93 to 1.7)</td>
<td>21 (10 to -179)</td>
</tr>
<tr>
<td>Twelve months</td>
<td>3</td>
<td>1196</td>
<td>1.01 (0.77 to 1.3)</td>
<td>493 (25 to -27)</td>
</tr>
</tbody>
</table>
Comment

This is smashing stuff in more ways than one. Interesting is that the points on the graph fall right on the line of equality. The authors make some comments on the blinding of acupuncture studies (difficult, on the overall quality of randomisation (generally poor), and on the fact that few studies tested for unequivocal smoking cessation. These factors would tend towards bias for acupuncture, which makes the profound negative even worse. The paper tells us that acupuncture for smoking is a poor buy.

We also have to beware thinking of placebo response here as people stopping smoking because of some feature of the treatment other than acupuncture. The reality is that the smoking cessation rates vary widely because the studies are small. They are not much different (early time point 0-54%, weighted average 27%; six months 0-34%, weighted average 14%) from the sorts of cessation rate with nicotine replacement (Bandolier 54). Here the average placebo rate for nicotine replacement was about 12% with at least six months of follow up.

People entering trials of smoking cessation want to stop smoking. Some of them succeed. With acupuncture, no more succeed. With nicotine replacement about 1 more will succeed for every 13 people who use it.

Reference:
1 AR White, H Rampes, E Ernst. Acupuncture for smoking cessation (Cochrane review). In Cochrane Library issue 1, 2000.

Book Reviews

Management of the Menopause. M Rees, DW Purdie. BMS Publications, Marlow. 1999. 78pp. ISBN 0 9536228 0 0. £10 plus postage from British Menopause Society, 36 West St, Marlow, Bucks SL7 2NB.

This is a useful monograph which seeks to make available an unbiased and non-promotional book on the menopause and its management. It does what it sets out do, and will be a useful read for many professionals, and for any woman approaching or experiencing the menopause. There is much to interest, and it is full of excellent common sense. It would be hard not to find something to appreciate.

It suffers from being a bit old-fashioned in its approach, in that it fails to inform as much as Bandolier would like. Some statements are unsupported by evidence. While there may be no evidence other than good clinical judgement (and there’s nothing wrong with that in the absence of good unbiased evidence), we don’t know that the absence of the evidence means the evidence is absent.

Quibbles apart, this is worth having on the shelves. Non-professionals will find this of interest also, so it might be valuable for patient bookshelves as well.

Clinical Evidence emerged from some ideas about creating an evidence formulary – a bit like the BNF with useful numbers in it. In that sense it really is an enormous undertaking, and its creators and editors are to be congratulated for even trying. It is impossible to do everything at once, so Clinical Evidence will grow and expand and get better.

It makes an excellent beginning. It tells you up front how it is put together, and contains a superb little glossary of technical terms to help you navigate through the knowledge. It tries to answer clinical questions. In this it really does seem to work very well. For instance, the section on skin diseases contains an item on head lice. Questions Bandolier was asked recently, about the evidence on nit combs or herbal or aromatherapy for head lice, were answered simply by reference to this. The questions were asked in Clinical Evidence, but no evidence was found.

The book tells you quite clearly that they try very hard to differentiate between evidence of lack of effectiveness, and lack of evidence. It does this well.

One way of telling whether a book like this is useful is to turn to a few subjects you know something about. If you do this, you will probably be pleasantly surprised. It does well. If there is any failing it is that the way results are described varies a bit between sections – some will give you odds ratios or absolute risk reductions. Others will give you NNTs. But most of us with a couple of neurones to rub together can hack this.

Bottom line to the editors is to keep up the good work. Bottom line for potential readers who are not lucky enough to get this free is to bite the bullet and buy it.


Point of care testing, including hospital hand-held and near patient devices, as well as home care, is big business: it will amount to over $5 billion in 2001. It is not just big business, but the generator of important sources of change in health services and their management. They also impact on patients, and in the environment where patients as customers demand more and better, the push to use point of care testing will only increase.

Ok, but what’s the technology, how do you manage it, and when is it effective and when not effective? Read this book and you will find answers to most of those questions, in detail and with a philosophical and critical background. The book is up to the minute, and covers a variety of clinical settings, from the home to intensive care. It is packed with practical information from informed people, and, while it is long at just under 600 pages, the pages just seem to fly by.

The only difficulty for those of us outside the USA may be obtaining it. But the American Association of Clinical Chemistry now has an Internet site so you can buy it online.