Who blows your whistle?

Winston Churchill never read newspapers. His wife used to read the newspaper for him and then give him a précis plus important cuttings. Clementine, adding wisdom, probably made a better job of converting information to knowledge than the software. It was a system that prevented him being deluged by paper and having the “frog in a jam jar” feeling.

David Halberstam’s brilliant book, ‘The Brightest and The Best’, has some marvellous insights into how institutions can stifle individuals. Individuals who did not agree with American Government policy on the Vietnam War were systematically sidelined. The information, the knowledge, which they produced to show that the policies were ineffective, was discarded in favour of the military bluff. The ability of individuals to speak out, even though we may think their views are crazy, is very precious.

Accepting change

It is by listening to whistleblowing that we can make progress through a process of constructive disagreement. Constructive disagreement is particularly important when we examine the evidence-base of what we do now or plan to do in the future. Change is often disagreeable, and it is surprising that in an area like healthcare where change is the rule rather than the exception, that we do not have more information on change management.

Where we do not believe

In Bandolier 42 we ran a Question Time to see if readers’ questions for which we could not find answers were answerable by others. About half the questions attracted responses of various sorts, which were passed on to the questioners, but there was insufficient evidence for a Bandolier article. One was more interesting, the use of glucosamine in arthritis.

It seemed at first that there was little evidence, and that perhaps it was negative. But we eventually found quite a number of papers (most were in the latest issue of the Cochrane Library), and so this issue has a brief review of what Bandolier was able to find. The bottom line is that it seems to work.

So, just like Hypericum for depression (Bandolier 31), an unconventional approach seems to be effective. But other conventional and unconventional approaches are not supported by evidence. For instance, we report another nail in the coffin of homeopathy this month. So headaches for believers and unbelievers alike, and difficult decisions for policymakers, patients, and their carers.

Glucosamine and arthritis

In Bandolier’s Question Time (Bandolier 42), Dr Proctor from Gainsborough asked what evidence there was that glucosamine was an effective treatment for arthritis. Given the amount of publicity it had been getting, we thought this was a good question. We had a number of responses, and decided to follow this up by trying to find papers and assess what evidence we could find.

Searching

This involved MEDLINE searching for articles on glucosamine, finding Internet pages which featured glucosamine, reference lists from retrieved articles, and references provided by Bandolier readers.

Results

Eight randomised trials [1-8] involving oral or intramuscular glucosamine were found and the details are in the large Table on page 3. Articles on intra-articular injection, or where the material used was not clearly defined as glucosamine were excluded, as were non-randomised case series. All those included examined oral and/or intramuscular glucosamine in patients with arthritis over periods of up to eight weeks. Most had well-described methods and six had quality scores of 3 or more on a five-point scale [9]. Oral doses of glucosamine sulphate were 1.5 grams a day, and intramuscular doses were 400 mg twice or three times a week.

Placebo-controlled trials

Five trials compared glucosamine with placebo. All showed statistical superiority of glucosamine. Four of these had dichotomous outcomes for calculating NNTs, which ranged from 1.7 to 6.3 in individual trials. Overall the NNT was 5.0 (3.5 to 8.9). This means that one of every five patients with
arthritis who are treated with glucosamine, one would have short term benefits in reduced pain and tenderness who would not have had if they had been given a placebo.

**Active-controlled trials**

Three trials compared glucosamine with NSAID (phenylbutazone or ibuprofen). There was no difference between ibuprofen (1.2 g/day) and oral glucosamine (1.5 g/day).

**Adverse effects**

Few adverse effects or study withdrawals were reported for glucosamine. They tended to occur less frequently with glucosamine than with NSAID. A large open study of 1208 arthritis patients taking oral glucosamine 1.5 g/day for 13 to 99 days [10] had 28 patients who stopped taking glucosamine because of adverse effects. Those adverse effects reported in more than 1% of patients were epigastric pain/tenderness, heartburn, diarrhoea and nausea.

**Comment**

*Bandolier* was surprised to find as many as eight randomised trials. While it is possible to criticise all of the trials to some extent, as a group they are no worse than others used to support commonly-used therapies. There is a consistent thrust of efficacy over placebo, and an inability to distinguish glucosamine from NSAID. But all trials were relatively short-term, and longer-term observations for adverse effects would be welcome.

The bottom line is that there is a body of evidence supporting the efficacy of oral and intramuscular glucosamine in arthritis.

---

**References:**

### RANDOMISED CONTROLLED TRIALS OF ORAL AND INTRAMUSCULAR GLUCOSAMINE

<table>
<thead>
<tr>
<th>Reference</th>
<th>Patients</th>
<th>Design</th>
<th>Drugs &amp; Doses</th>
<th>Outcomes</th>
<th>Results</th>
<th>Quality score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mund-Hoym, 1980</td>
<td>80 patients with &quot;vertebral syndrome&quot;. Mean age 58 years</td>
<td>Randomised, parallel-group</td>
<td>3 injections of 400 mg glucosamine sulphate a week, plus oral glucosamine 250 mg 2/3 times a day on non-injection days (40 patients) Daily injections of 600 mg phenylbutazone (40 patients) All IM injections</td>
<td>Number of days to clinical improvement. Global good/bad</td>
<td>Glucosamine 32 days (± 1.3, range 21 - 48). Phenylbutazone 46 days (± 1.7, range 27 - 80). Good/bad outcome 34/6 for glucosamine, 29/11 for phenylbutazone</td>
<td></td>
</tr>
<tr>
<td>Pujalte et al, 1980</td>
<td>24 patients with osteoarthritis of the knee. Mean age 60 years</td>
<td>Randomised, double-blind, parallel-group, 6-8 weeks</td>
<td>500 mg glucosamine sulphate, three times daily or identical placebo. No other analgesics allowed</td>
<td>Articular pain, swelling, movement using categorical scale (patient and doctor)</td>
<td>Glucosamine significantly (p&lt;0.01) better than placebo for pain, tenderness and swelling. Time for clinical improvement 14 days for glucosamine cf 40 days for placebo. No AE with glucosamine. Glucosamine significantly (p&lt;0.01) better than placebo for pain and function restriction at 21 days. Overall symptom score better with glucosamine. No difference in walking times, though better. No AE.</td>
<td></td>
</tr>
<tr>
<td>Crollie &amp; D’Este, 1981</td>
<td>30 in-patients with chronic osteoarthritis. Mean age 72 years</td>
<td>Randomised, double-blind, parallel-group, 21 days</td>
<td>400 mg intramuscular glucosamine sulphate for 7 days, followed by 500 mg orally three times a day for 14 days. IM piperazine/chlorbutanol for 14 days, followed by oral placebo for 14 days.</td>
<td>Pain at rest or movement, categorical scale. Walking time over 20 metres.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D’Ambrosio et al, 1981</td>
<td>30 in-patients with chronic degenerative osteoarthritis. Mean age 75 years. No steroids or NSAIDs for 2 weeks before trial.</td>
<td>Randomised, open, parallel-group, 21 days</td>
<td>400 mg intramuscular glucosamine sulphate for 7 days, followed by 500 mg orally three times a day for 14 days. IM piperazine/chlorbutanol for 14 days, followed by oral placebo for 14 days.</td>
<td>Pain at rest or movement, categorical scale.</td>
<td>Glucosamine significantly (p&lt;0.01) better than placebo for symptom score at 21 days. (pain at rest and movement appear significantly improved, though no stat test done). No AE.</td>
<td></td>
</tr>
<tr>
<td>Vaz, 1982</td>
<td>40 out-patients with unilateral osteoarthritis of the knee without complications. Mean age 58 years.</td>
<td>Randomised, double-blind, parallel-group, 8 weeks</td>
<td>1.5 g/day of glucosamine sulphate or 1.2 g ibuprofen.</td>
<td>Pain, categorical scale.</td>
<td>Ibuprofen significantly better than glucosamine at 2 weeks, glucosamine significantly better than ibuprofen at 8 weeks. AE (mild) reported by 2 on glucosamine and 5 on ibuprofen. Overall efficacy (doctor) good 8/18 glucosamine and 3/20 ibuprofen.</td>
<td></td>
</tr>
<tr>
<td>Rovati, 1992, Study 1</td>
<td>252 out-patients with gonarthrosis.</td>
<td>Randomised, double-blind, parallel-group, 4 weeks</td>
<td>1.5 g/day of glucosamine sulphate or placebo.</td>
<td>Lesquesne index, responders/non-responders</td>
<td>Glucosamine significantly (p&lt;0.05) better than placebo for symptom score at 4 weeks. Responders 66/126 glucosamine, 46/126 placebo. Minor AE 8/126 glucosamine, 13/126 placebo.</td>
<td></td>
</tr>
<tr>
<td>Reichelt et al, 1994 (duplicated in Rovati, 1992, Study 2)</td>
<td>155 out-patients with gonarthrosis.</td>
<td>Randomised, double-blind, parallel-group, 6 weeks</td>
<td>Intramuscular glucosamine sulphate 400 mg or placebo twice a week.</td>
<td>Lesquesne index, responders/non-responders</td>
<td>Glucosamine significantly (p&lt;0.04) better than placebo for symptom score at 4 weeks. Responders 40/79 glucosamine, 23/76 placebo. Minor AE 5/79 glucosamine, 3/76 placebo.</td>
<td></td>
</tr>
</tbody>
</table>

**CONSTIPATION**

We all know what it is, but it is hard to define. Bowel movement patterns show that 90% of people in Western countries have between three bowel movements a day to three per week. Constipation is often defined as fewer than three bowel movements a week, though symptoms like straining, passing hard stools and inability to defecate when desired, together with abdominal pain also form part of a diagnosis.

The use of laxatives, both prescribed and non-prescribed, is common, and constipation is thought to be common, especially in older people, women, and people with poor diets. So one might think that there would be a wealth of information from clinical trials on the effectiveness of laxative agents - that we would know that they work, and how well they work.

Bandolier thought so too. So it was a surprise to see from a systematic review [1] that we don’t have as much information as we thought, or would like.

**Review**

This is a good review, with great searching (though some might quibble about the exclusion of reports not in English), and an interesting description about a large number of trials which did not pass muster for various reasons. It makes a good read, though the journal is not readily available and Bandolier had to obtain it through the British Library.

Trials were included if they looked at treatment of constipation in adults for at least two weeks (minimum of one week on treatment or control).

**Results**

They found 36 trials with 1,815 subjects, of whom 40% were over 60 years, and 70% of whom were women. Many of the trials had poor design. There were 25 different laxative or dietary fibre therapies. Twenty trials compared laxative with placebo or regular diet, and 16 were comparisons with other laxatives.

**Bowel movements**

The number of bowel movements per week in controls ranged from a mean (or median) of 1.5 to 7.1. Only six of 16 reports had mean bowel movements per week below 3. Laxatives increased the number of bowel movements (Figure) from a control mean (weighted by number of patients) of 3.5 per week to 5.0 per week. The increase was 1.4 bowel movements per week (95% confidence interval 1.1 to 1.8).

Bulk laxatives (six trials) gave an average weighted mean increase of 1.4 (0.6 to 2.2), and other agents (seven trials) gave a weighted mean increase of 1.5 (1.1 to 1.8).

Overall symptom improvement was reported as being significantly better for laxatives than control in nine of eleven trials which measured them.

Results from the trials which compared different laxatives were largely uninterpretable.

**Comment**

Laxatives work. Perhaps there’s nothing new in that. What is disappointing is that there is so little evidence to allow us to see which is best. What we can say is that for adult patients with chronic constipation, bran or bulk laxatives work as well as anything. The amount of bran used in the trials ranged from 0.5 gram to 24 grams a day, equivalent to from a quarter of a serving of fruit, vegetables or cereal a day to 12 servings a day. Advising adults with chronic constipation to eat more fruit and vegetables and have some bran seems to be the best evidence available.

References:
More on BPH

Treatment of benign prostatic hypertrophy has been further informed by a meta-analysis and a large randomised controlled trial (RCT). The meta-analysis looked at results after one year on finasteride or placebo, and the RCT examined effects after two years of treatment. Bandolier has a quibble with both of these reports. Each has lovely statistics and large numbers, but both lack a clear statement of what constitutes clinical improvement in men with BPH, either in terms of peak urine flow rate, or in terms of symptom scores. And without that, and with just means, no NNTs can be calculated from the published data, which is disappointing considering the quality of the studies.

Meta-analysis

This examined one-year results of six RCTs comparing finasteride and placebo [1]. There were 2601 men randomised between finasteride 5 mg or placebo daily for one year. Prostate volume was measured by trans-rectal ultrasound (TRUS) or MRI.

The bottom line was that finasteride is only effective in men with prostate volumes of more than 40 mL. The Figure shows data from a single patient analysis for peak urine flow rate, though a similar picture emerges for symptom scoring. The conclusion was that men with small prostates may not be suitable candidates for finasteride therapy.

RCT - two year outcomes

The RCT enrolled 707 men with moderate symptoms of BPH and treated them with 5 mg finasteride or placebo daily for two years.

The main results were that men on placebo had worsening symptom scores and urine flow rates in the second year of treatment, whereas men on finasteride maintained the benefits seen in the first year of treatment. There was sufficient information to calculate some NNT values (Table). For instance, prostate volumes increased in 56% of men on placebo, but only 16% of men on finasteride. The NNT of 2.5 indicates that finasteride has to be given to five men for two years to prevent an increase in prostate volume of more than 1 mL in two, in whom this would not have been prevented with placebo.

Finasteride versus placebo for BPH: two-year outcomes

<table>
<thead>
<tr>
<th>Outcome</th>
<th>NNT (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevent prostate volume increase &gt;1 mL</td>
<td>2.5 (2.1 to 3.2)</td>
</tr>
<tr>
<td>Prevent one patient developing urinary retention</td>
<td>32 (18 to 136)</td>
</tr>
<tr>
<td>Prevent one patient having prostate surgery</td>
<td>39 (23 to 111)</td>
</tr>
<tr>
<td>Prevent one patient discontinuing because of insufficient response</td>
<td>39 (17 to no difference)</td>
</tr>
<tr>
<td>Cause one case of sexual dysfunction</td>
<td>11 (6.9 to 23)</td>
</tr>
</tbody>
</table>

One patient in 30 to 40 would stop the drug because of insufficient response, or because of urinary retention or surgery. For finasteride to cause one case of sexual dysfunction had an NNT of 11.

Comment

The accumulating information on finasteride treatment of BPH is giving more clear indications on which men can benefit (those with prostate volumes >40 mL) and that the benefits seen in one-year studies continue at least to two years. The lack of information to calculate NNTs for clinically relevant outcomes (and what they might be) still eludes us.

References:

Meta-analysis of six RCTs, effect of finasteride and placebo on peak urine flow rate at one year

![Graph showing peak urine flow rate change from baseline (mL) over prostate volume by TRUS or MRI (mL).](image)
**Signs and Symptoms Predict Thyroid Disease**

On the face of it this headline has all the impact of “dog bites man”. But one of the problems that clinical laboratories face, as they have for years, is the overwhelming tide of tests. When all they could offer was basic metabolic rate (and those old enough will remember how difficult they were to do), patients had full clinical workups before the test was done. A simple blood test is just that, simple, so anyone with any possibility of disease is investigated. All laboratories have a library of what they would regard as stupid reasons for requesting thyroid function tests - ingrowing toenail is one Bandolier remembers.

So some simple words from 20 years ago about the relationship between clinical signs and symptoms and the incidence of thyroid disease [1] are still relevant.

**Study**

Five-hundred consecutive inpatients and outpatients (those with known thyroid disease being excluded) to Flinders Medical Centre had case notes examined for:

- thyroid function test (TFT) results.
- clinical signs and symptoms noted by the clinician during the consultation from which the referral for TFT was made.
- subsequent clinical diagnosis of thyroid status.

The degree of clinical suspicion from signs and symptoms from a generally accepted (see box) list was correlated with final outcome.

**Results**

Of the 500 patients, 21 (4.2%) were found to have thyroid dysfunction needing treatment.

In those patients with five or more clinical signs or symptoms in whom the degree of clinical suspicion was high, the majority (18, 78%) had thyroid disorder needing treatment. Using final diagnosis as the gold standard, the likelihood ratio for having a high degree of clinical suspicion was 82. This meant that from a pre-test probability of 4.2% the post-test probability was over 80%.

In those with high or intermediate degree of clinical suspicion, 19 (33%) had thyroid disorder needing treatment. Using final diagnosis as the gold standard, the likelihood ratio for having a high or intermediate degree of clinical suspicion was 11. This meant that from a pre-test probability of 4.2% the post-test probability approached 40%.

In those with low degree of clinical suspicion, 2 of 442 patients (0.45%) had thyroid disorder needing treatment. Using final diagnosis as the gold standard, the likelihood ratio for having a low degree of clinical suspicion was 0.1 (calculated as the likelihood ratio of a negative test). This meant that from a pre-test probability of 4.2% the post-test probability was less than 1%.

<table>
<thead>
<tr>
<th>Degree of suspicion</th>
<th>Number of patients</th>
<th>Number with thyroid disease</th>
<th>Percent with thyroid disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>23</td>
<td>18</td>
<td>78</td>
</tr>
<tr>
<td>Intermediate</td>
<td>35</td>
<td>1</td>
<td>2.9</td>
</tr>
<tr>
<td>Low</td>
<td>442</td>
<td>2</td>
<td>0.45</td>
</tr>
<tr>
<td>Total</td>
<td>500</td>
<td>21</td>
<td>4.2</td>
</tr>
</tbody>
</table>
Likelihood ratio nomogram
signs & symptoms
of thyroid disease

Pre-test probability

Post-test probability

Likelihood ratio

5+ symptoms

3+ symptoms

0, 1 or 2 symptoms

Bandolier has always found these words comparing clinical examination and exercise testing compelling, but has struggled to find examples. Thyroid function testing appears to be one. Of course, these are “old” data, but Australian patients being examined in 1977 won’t be that different from people in the UK in 1997.

Since the incidence of thyroid disease in the general population is the order of 1% or below, some selection has already gone on to create a 4.2% incidence in the population in the paper. But that serves only to emphasise the very low yield of thyroid disease (0.45%) in those patients in whom there was a low clinical suspicion. It prompts the question, why do a TFT? But post-test probabilities of more than 30% for at least 3 signs and symptoms, and of more than 80% for at least 5 signs and symptoms would seem to make TFTs worthwhile in confirming the diagnosis.

The authors discuss the “heavy economic and logistic burden of biochemical screening” that biochemical screening was causing. Plus ça change.

References:
1 GH White, RN Walmsley. can the initial clinical assessment of thyroid function be improved? Lancet 1978 ii: 933-5.

Study

During a 10-month period in 1982/3, patients admitted to San Francisco General Hospital coronary care unit were eligible for entry. Of these, 393 entered the trial and 57 did not want to participate after being fully informed about the nature of the project.

Intercessionary prayer was provided by “born again” Christians of several denominations. After randomisation (by computer-generated list), patients in the prayer group were prayed for by between three and seven intercessors. The intercessors were given the first name, diagnosis and general condition of the patient, with pertinent updates. Prayer took place outside the hospital daily until discharge. Intercessors prayed for a rapid recovery, and for prevention of complications and death, as well as anything else they wanted to add to the prayer.

Patients had no idea whether or not they were being prayed for. Additional prayers in either group by, for instance, family members, was not controlled for. Data on patients’ condition, complications and outcome was collected blind.

Results

There were no differences at entry between patients for any demographic variable, primary cardiac diagnosis, or noncardiac illness or complication.

Intercessionary prayer was without effect on days spent in the coronary care unit, or days in hospital, or number of discharge medications. There were 26 new problems, diagnoses or therapeutic events monitored after entry into the trial; 107 events occurred in 192 patients being prayed for and 175 occurred in 210 control patients. Six events occurred significantly less frequently with prayer - congestive heart failure, use of diuretics, cardiac arrest, pneumonia, antibiotic use and intubation or ventilation.

Comment

“Look at the relative size of the likelihood ratios for a brief, immediate, relatively cheap history and a much longer, delayed, and relatively expensive exercise electrocardiogram. There is no contest. Likelihood ratios for key points in the history and physical examination, both for this and for most other target disorders, are mammoth and dwarf those derived from most excursions through high technology.”[2]

O LD CURIOSITY SHOP

The power of prayer

Many people of different faiths believe that prayer can have beneficial effects on their own or others’ health. Testing the power of prayer is not easy, but a randomised, double-blind trial with large numbers demands some attention.

References:
1 GH White, RN Walmsley. can the initial clinical assessment of thyroid function be improved? Lancet 1978 ii: 933-5.
The clinical course of patients was scored as good, intermediate, or bad according to a scoring system. Good outcomes were more frequent in patients who were being prayed for (163/192; 85%) than in those who were not prayed for (147/201; 73%). This generated a relative benefit of 1.16 (1.05 - 1.29) and a NNT of 8.5 (5.1 - 26).

Comment

People will be able to see what they want from this trial. It was a properly randomised, double-blind trial. It had statistically significant outcomes in favour of prayer having a beneficial effect in this patient group. There were no outcomes for which the control group did better than those being prayed for, and though the effects are not great, they all go in one direction, that prayer is effective.

Doubters might point out that there may have been an element of data-dredging, because there was no prior statement of what outcomes were going to be looked at, and so we have to look less favourably on the result. They may also point out that some of the events where statistical significance was found were fairly rare, occurring in only a few percent of patients in both groups, so random chance may play a part.

The lesson is that a single trial is just that, one observation. If the effect was massive, and the trial huge, and there was an agreed and understood mechanism, then perhaps taking results from a single trial may be OK. Where these conditions are not met, then caution rules. The fact that a Cochrane review group is summarising all the literature on the effects of prayer is welcome.

Reference:

Effect of prayer on hospital course

<table>
<thead>
<tr>
<th>Percent</th>
<th>Good</th>
<th>Intermediate</th>
<th>Bad</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Prayer Control

DIARY NOTE

The Bandolier conference on Chlamydia had had to be postponed. It will now take place on March 27, 1998, at the Wellcome Institute in London.

Details can be obtained from Eileen Neail, preferably by fax on 01865 226978.