**Incidence/Prevalence**

One overview of observational studies published in English found that the prevalence of GAD among adults in the community is 1.5–3.0%. It found that 3–5% of adults have had GAD in the past year and 4–7% have had GAD during their lifetime. The US National Comorbidity Survey found that over 90% of people diagnosed with GAD had a comorbid diagnosis, including dysthymia (22%), depression (39–69%), somatisation, other anxiety disorders, bipolar disorder, or substance abuse. The Harvard Brown Anxiety Research Program also found that only 30/180 (17%) people had GAD alone. Subgroup analysis suggested that 46/122 (38%) of people with GAD had comorbid personality disorder. A systematic review of the comorbidity of eating disorders and anxiety disorders (search date 2001, two observational studies, 55 people) found a lifetime prevalence of GAD among people with anorexia nervosa of 24% in one study and 31% in the other. The lifetime prevalence of GAD in the control group of one of the studies (44 people) was 2%. The reliability of the measures used to diagnose GAD in epidemiological studies is unsatisfactory. One US study, with explicit diagnostic criteria (DSM-III-R), estimated that 5% of people will develop GAD at some time during their life. A recent cohort study of people with depressive and anxiety disorders found that 49% of people initially diagnosed with GAD retained this diagnosis over two years. The incidence of GAD in men is only half the incidence in women and is lower in older people. A non-systematic review (20 observational studies in younger and older adults) suggested that autonomic arousal to stressful tasks is decreased in older people, and that older people become accustomed to stressful tasks more quickly than younger people.

**Aetiology/Risk Factors**

Generalised anxiety disorder is believed to be associated with an increase in the number of minor stressors, independent of demographic factors, but this finding is also common in people with other diagnoses in the clinical population. One non-systematic review (five case control studies) of psychological sequelae to civilian trauma found that rates of GAD reported in four of the five studies were significantly increased compared with a control population (rate ratio 3.3, 95% CI 2.0 to 5.5). One systematic review (search date 1997) of cross sectional studies found that bullying (or peer victimisation) was associated with a significant increase in the incidence of GAD (effect size 0.21). Genetic factors are also implicated. One systematic review (search date not reported, two family studies, 45 index cases, 225 first degree relatives) found a significant association between GAD in the index cases and their first degree relatives (OR 6.1, 95% CI 2.5 to 14.9). The review also identified three twin studies (13 305 people), which estimated that 32% (95% CI 24% to 39%) of the variance to liability to GAD was explained by genetic factors.

**Prognosis**

One systematic review found that 25% of adults with GAD will be in full remission after two years, and 38% will have a remission after five years. The Harvard-Brown anxiety research program reported five year follow up of 167 people with GAD. In this period, the weighted probability for full remission was 38% and for at least partial remission was 47%; the probability of relapse from full remission was 27% and relapse from partial remission was 39%.

**What are the Effects of Treatments?**

**Likely to be beneficial**

- **Buspirone**
  Randomised controlled trials (RCTs) have found that buspirone improves symptoms compared with placebo over 4–9 weeks. RCTs found no significant difference in symptoms over 6–8 weeks between buspirone and antidepressants, diazepam, or hydroxyzine, but the studies may have lacked power to detect clinically important differences among treatments.

- **Certain antidepressants** (imipramine, opipramol, paroxetine, and venlafaxine)
  Randomised controlled trials have found that antidepressants (imipramine, opipramol, paroxetine, and venlafaxine) improve symptoms over 4–28 weeks compared with placebo. RCTs found no significant difference among these antidepressants or between antidepressants and benzodiazepines or buspirone. RCTs and observational studies have found that antidepressants are associated with sedation, dizziness, nausea, falls, and sexual dysfunction.

- **Cognitive behavioural therapy**
  Two systematic reviews and two subsequent RCTs have found that cognitive behavioural therapy (using a combination...
of interventions, such as exposure, relaxation, and cognitive restructuring) improves anxiety and depression over 4–12 weeks compared with waiting list control, anxiety management alone, relaxation alone, or non-directive psychotherapy. Three subsequent RCTs, two in people aged ≥60 years, found no significant difference in symptoms at 13 weeks, six months, or 24 months between cognitive therapy and applied relaxation.

Hydroxyzine
Three RCTs comparing hydroxyzine versus placebo found different results. Two RCTs found that, compared with placebo, hydroxyzine improved symptoms of anxiety at four or 12 weeks, but a third RCT found no significant difference in the proportion of people with improved symptoms of anxiety at five weeks. One of the RCTs found that hydroxyzine increased somnolence and headaches compared with placebo. One RCT found no significant difference between hydroxyzine and bupropion in the proportion of people who responded after six weeks. Another RCT found no significant difference between hydroxyzine and buspirone in the proportion of people who responded after four weeks.

Trade off between benefits and harms
Benzodiazepines
One systematic review and one subsequent RCT found that benzodiazepines reduced symptoms over 2–9 weeks compared with placebo. RCTs found no significant difference in symptoms over 3–8 weeks between alprazolam and bromazepam or mexital, or between benzodiazepines and buspirone, hydroxyzine, abecarnil, or antidepressants. RCTs and observational studies found that benzodiazepines increased the risk of dependence, sedation, industrial accidents, and road traffic accidents and that, if used in late pregnancy or while breast feeding, benzodiazepines may cause adverse effects in neonates. RCTs found no significant difference in symptoms over 3–8 weeks between alprazolam and bromazepam or mexital, or between benzodiazepines and buspirone, hydroxyzine, abecarnil, or antidepressants. One systematic review of poor quality RCTs provided insufficient evidence to assess long term treatment with benzodiazepines.

Kava
One systematic review in people with anxiety disorders, including generalised anxiety disorder, found that kava reduced symptoms of anxiety over 1–24 weeks compared with placebo. It is unclear whether results of the review are generalisable to people with generalised anxiety disorder. Observational evidence suggests that kava may be associated with hepatotoxicity.

Trifluoperazine
One large RCT found that trifluoperazine reduced anxiety after four weeks compared with placebo, but caused more drowsiness, extrapyramidal reactions, and other movement disorders.

Unknown effectiveness
Abecearnil
One RCT found limited evidence that low dose abecarnil improved symptoms compared with placebo. Another RCT found no significant difference in symptoms at six weeks between abecarnil and placebo or diazepam. Both RCTs found that abecarnil increased drowsiness compared with placebo.

Applied relaxation
We found no RCTs comparing applied relaxation versus placebo or no treatment. Three RCTs found no significant difference in symptoms at 13 weeks, six months, or 24 months between applied relaxation and cognitive behavioural therapy.

β blockers
We found no RCTs on the effects of β blockers in people with generalised anxiety disorder.

REFERENCES
17 Hawker DSJ, Boulton MJ. Twenty years’ research on peer victimisation and psychosocial maladjustment: a meta-analytic review of cross-sectional studies. J Child Psychol Psychiat 2000;41:441–5. Search date 1997; primary sources PsychLit, Social Science Citation Index, OCLC FirstSearch, hand searches of relevant journals, bibliographies, reviews, reference lists of relevant articles, and book chapters, and personal contact with authors.
Meta-analyses and megatrials: neither is the infallible, universal standard

Nowadays, most would agree that we need evidence from randomised control trials (RCTs) to evaluate the effectiveness of a health intervention. It used to be that we did not have enough RCTs in mental health; the irony today is that at times it seems we have too many of them, especially when they draw conflicting conclusions.

A natural solution is to seek “stronger” evidence. Meta-analysis might provide that evidence but, alas, meta-analyses sometimes do not agree among themselves either. Another possible solution is a bigger and better trial, a megatrial (also known as the large, simple trial). Unfortunately megatrials and meta-analyses do not always agree either: one group has claimed that—taking megatrials as the gold standard—meta-analyses drew wrong conclusions 35% of the time; another group estimated the degree of disagreement to be between 10% and 23%. Megatrials sometimes do not agree with each other either, and discrepancies among megatrials are just as large as those between meta-analyses and megatrials.

These discrepancies reinforce a conclusion that the days of dogmatic advocacy of the methodological hierarchy of evidence are over. Here, I will take three examples to illustrate that we will always need “good common sense”, coupled with content expertise and an understanding of methodology, to weigh the available evidence relevant to a mental health problem.

<table>
<thead>
<tr>
<th>Table</th>
<th>Strengths and weaknesses of meta-analyses and megatrials</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Definition</td>
</tr>
<tr>
<td></td>
<td>Apply to both meta-analyses and megatrials</td>
</tr>
<tr>
<td></td>
<td>Megatrial Very large randomised controlled trials, usually recruiting thousands of subjects and usually multicentred. Recruitment criteria are very broad, protocols are maximally simplified, and endpoints are unambiguous, such as death. Also often referred to as a “large, simple trial”. Typical examples are seen in cardiovascular medicine.</td>
</tr>
<tr>
<td></td>
<td>Meta-analysis Combination of data from several independently performed single or multicentre trials with the purpose of assessing effects on endpoints for which the individual trials are usually non-informative due to lack of statistical power.</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Well conducted systematic reviews including megatials usually offer the best guide to overall treatment effect. For example, in the case of risperidone versus typical antipsychotics for schizophrenia, a very large multinational, multicentre RCT (n = 1362) found no statistically significant difference between these two drugs (RR of no response = 0.94; 95% CI 0.79 to 1.11). A subsequent Cochrane review that included an additional 1006 subjects did show, in contrast, a significant and important random effects RR of 0.84 (95% CI 0.76 to 0.92) in favour of risperidone. There was no indication of heterogeneity across trials (p = 0.63). It appears that one of the largest trials to date in mental health was still underpowered to detect a small yet important difference.

When available studies for meta-analysis are limited in number, sample size, or quality of methodology, we are in a more difficult position. Another Cochrane review concluded that lithium therapy is an efficacious maintenance treatment for bipolar disorder. Combining three studies (total n = 412), this review found a statistically significant and clinically meaningful reduction in relapse for patients with bipolar disorder on lithium compared to placebo (random effects RR = 0.60, 95% CI 0.41 to 0.87). Heterogeneity among the included RCTs was not statistically significant (p = 0.13) but substantive ($I^2 = 51.6\%$). Although two older studies found lithium to be superior to placebo, the most recent study failed to find a statistically significant difference between the two arms (RR = 0.71, 95% CI 0.39 to 1.31). One reasonable conclusion was that the latest study was underpowered and was in fact in concordance with previous studies. Considerable debate ensued after publication of this pivotal study and there was ongoing debate on the methodological adequacy of the older trials with lithium. The superficial interpretation of the more recent study as "negative" seemed to support claims against accepted wisdom in modern psychiatry. This clinical and scientific chagrin abated somewhat when the same group of researchers published a similarly planned maintenance RCT and found a significant reduction in relapse on lithium in comparison with placebo. Closer reading of their report reveals, however, that lithium reduced relapse over 12 months only at the expense of increasing dropouts due to adverse events; survival on the medication without relapse or dropout was no different on lithium or on placebo. Only 22% and 16%, respectively, of those starting on lithium or placebo remained on the same drug without relapse until the study termination up to 18 months. The value of lithium appears small at best.

When a systematic review is of inferior quality, we are in an even more difficult position. A meta-analysis of alprazolam for anxiety disorders involving 8878 randomised patients claimed to have confirmed its efficacy. Alprazolam may indeed be better than placebo in reducing panic and associated anxiety over 8–12 weeks but we need no more than a well designed, well analysed study of 154 patients to convincingly disqualify alprazolam as drug of choice for anxiety disorders. Aprazolam alone was not as good as exposure therapy alone for the acute phase of treatment, and the addition of alprazolam to exposure therapy resulted in even worse outcomes at follow up than exposure alone.

Having observed these illustrative cases and having appreciated that a thorough critical reading of a comprehensive meta-analysis is a formidable task, we in the Department of Psychiatry at Nagoya City University tend to examine meta-analysis as a navigator for sound evidence on a clinical topic. Looking at the whole map of available trials in the metaview of the Cochrane Library, we often choose to critically appraise and learn from the best—the largest, the most recent, the best known, the closest to the overall mean, whatever—trial in detail. We find that such practice often brings more insight to the bedside the next day than critically appraising the meta-analysis itself.

The strengths and weaknesses of meta-analyses and megatials are shown in table 1. We can never arrive at infallible truth because, firstly, that is simply not the nature of scientific knowledge and, secondly, in clinical medicine we are dealing with complex, ever changing units of analysis that are people with illnesses.

TOSHI FURUKAWA
Professor of Psychiatry, Evidence-Based Psychiatry Center (www.ebcenter.com), Nagoya City University, JAPAN; furukawa@med.nagoya-cu.ac.jp

Other articles noted

The journals that are reviewed and the criteria for selecting articles from these journals for inclusion in Evidence-Based Mental Health are set out in the purpose and procedure in each issue. These articles met inclusion criteria for Evidence-Based Mental Health, but were not abstracted due to lack of space.

**THERAPEUTICS**


**AETIOLOGY**


**PROGNOSIS**