A systematic meta-review of hypnosis as an empirically supported treatment for pain

Russell MF Hawkins

School of Psychology, University of South Australia and Flinders Medical Centre Pain Management Unit, Adelaide, South Australia

A newly developed methodological technique (the systematic meta-review) was applied to determine whether hypnosis is an empirically supported treatment for pain. This involved, initially, a systematic search of the published literature for review studies. These reviews were then subjected to a validated quality scale. There was sufficient evidence of good quality to allow the conclusion that hypnosis does have demonstrable efficacy in the treatment of pain. The only meta-analysis of hypnotically induced analgesia showed that ‘the average participant treated with hypnosis demonstrated greater analgesic response than 75% of participants in standard and no-treatment control groups’ (p. 143). The hypothesis that poor-quality reviews are more likely to produce positive conclusions was not supported. A citation database of all reviews has been assembled and can be extended with time.

Introduction

In the hypnosis literature, as well as in medicine and psychology more generally, the importance of adequate evidence of outcome has become prominent. This is exemplified by a special issue of the International Journal of Clinical and Experimental Hypnosis, published in April, 2000. This issue consists of commissioned articles intended to address the clinical efficacy literature according to the criteria developed by the American Psychological Association. These quality control standards were intended to offer a high degree of confidence in the validity of any conclusions reached. In the preamble to the special edition, the Editor noted that ‘hypnosis may be one of the most thoroughly researched forms of psychotherapeutic intervention, with more than 7,000 publications since 1966, in more than 150 different general medicine, psychological and interdisciplinary journals’ (p. 108).2

Montgomery et al.1 suggested that, even with the range of treatment contexts for which hypnosis is used, it is best known as a pain management technique. In a review of psychological interventions for chronic pain published about 20 years ago, Turner and Chapman3 remained sceptical about hypnosis, reporting that they were able to find no controlled studies comparing hypnosis with a credible placebo for the control of chronic pain. They concluded that ‘the clinical research in this area is sparse, appallingly poor, and has failed to convincingly demonstrate that hypnosis has more than a placebo effect in relieving chronic pain’ (p. 30). The present review will concentrate on studies published since that time in order to determine whether any alternative conclusion is now justified.

It is a remarkable experience to witness a person undergo surgery without an anaesthetic,
or to observe the lack of suffering in a woman giving birth, a burn patient undergoing debridement without pain, or a child tolerating well a bone marrow aspiration or lumbar puncture, after the use of hypnotic analgesia. On the basis of these impressive demonstrations, it is clear that hypnosis can be a powerful analgesic. No matter how compelling specific cases may be, this sort of evidence remains unsatisfactory and the usual rules of good science must be applied to the phenomenon before efficacy can be regarded as having been established. Anecdotal examples allow nothing to be said about the potential for the generalization of outcome to other cases.

In theory, it should be a relatively simple matter to determine whether particular clinical treatments are efficacious. The interested clinician may consult specialist journals and expect to find accounts of studies designed to demonstrate treatment outcome, read these accounts, and come away with a clear idea about whether or not a particular treatment was warranted in his or her clinical practice. In fact, the process of trying to quantify efficacy is fraught with difficulty because the quality of the evidence available for consultation is variable and often poor, and studies addressing the same therapeutic problem sometimes produce conflicting results.

The project of which this article forms a part includes an attempt to review all significant publications in the area of hypnosis and pain. The present work will concentrate on review articles, in which authors have already attempted to aggregate the available evidence about efficacy.

### Which evaluation standards could be used?

**Assessment of primary studies**

In order to draw conclusions about the benefits of hypnosis, an early task is to decide on the standards to be used for evaluation.

Relatively recent attempts to take seriously the quality of the evidence have been the cornerstone of the evidence-based medicine (EBM) movement, which developed from clinical epidemiology and medical informatics. A current definition of EBM is that it is ‘the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients’ (p. 2).

McQuay and Moore have suggested that EBM is actually what most clinicians have been trying to practise all of their working lives, but what is new is that there is an increasing number of well-conducted randomized controlled trials and systematic reviews, which can now be used to inform decisions about treatment. In the area of pain relief research there certainly has been an increasing volume of randomized controlled trial (RCT) publications over time. In five-year periods, the number of RCTs published in the journal *Pain* increased from 22 (1975–1979), to 42 (1980–1984), to 78 (1985–1989) and to 104 (1990–1994).

Although the RCT is known to be the most reliable way to estimate the effect of an intervention, even RCTs vary in their quality. For this reason a method for assessing the quality of published reports is important. This need prompted a search for measures of quality that could be utilized to assess the hypnosis and pain literature.

A number of quality measures for primary studies are available. Moher et al. provided an annotated bibliography of scales and check-lists designed to assess the quality of RCTs but Jadad et al. commented that, of the 25 scales identified in Moher et al.’s report, ‘only one had been developed following established methodological procedures’ (p. 2). A common weakness was that the various scales tended not to have been adequately assessed for validity. Jadad et al. developed a quality scale, which was based on established methodological procedures. After initially considering 49 items, the scale, which was redescribed by McQuay and Moore, eventually consisted of just the following three questions:

1) Was the study described as randomized?
2) Was the study described as double blind?
3) Was there a description of withdrawals and drop-outs?

The importance of each of these criteria will be considered briefly. In assessing a study using this approach, additional points are earnable or deductible according to the quality of the randomization and blinding effort, leading to a final score with a maximum of 5 points.
Randomization

The aim of randomization is simply that patients in a RCT should have the same probability of receiving any of the interventions being compared.

Randomization abolishes selection bias because it prevents investigators from influencing who has which intervention (methods of allocation based on alternation, date of birth, or hospital record number cannot be regarded as random).

Blinding

Blinding is an attempt to minimize or eliminate rater bias. Ideally, according to McQuay and Moore, rater bias would be minimized by blinding the person receiving the intervention, the individuals administering it, the investigator measuring the outcome, and the analyst. This extreme approach is seldom achieved; more commonly ‘double blinding’ is used, when the study participant and those making the observations are blinded. If double blinding has been successful, neither the patient nor the therapist should be able to detect at better than chance levels which treatment the patient has received.

Withdrawals and drop-outs

Bias may still occur after adequate randomization and blinding have been achieved if a systematic form of bias causes greater attrition (withdrawal and drop-outs) in one group as compared with another. For this reason it is important that investigators should report the incidence of attrition. The Journal of Evidence-Based Medicine will not publish trials with a less than 80% follow-up; that is, in which more than 20% of the sample dropped out.

Underlying the inclusion of questions about randomization and blinding in the McQuay and Moore quality measure is a finding that, when randomization was inadequate, there was an average 41% overestimation of the treatment effect and when there was no double blinding there was an average overestimation of treatment effect of 17%. Small trials also resulted in a 30% overestimation effect.

The McQuay and Moore quality measure seems well suited to certain types of interventions for pain, particularly for the assessment of drugs, where blinding is often easy to achieve. Psychological therapies are much more difficult to blind (although this has been attempted) and it is difficult to imagine any feasible way of producing a truly double blind study of hypnosis. An implication from this situation is that the lack of blinding means that studies of hypnotic intervention can never be judged as being of top quality. In principle this may be logical but in practice it is not helpful in terms of making relative quality judgements about existing hypnosis publications.

It is not feasible to blind either participants or investigators to the use of hypnosis, so it is necessary to consider alternatives to the McQuay and Moore quality measure that could still safeguard quality without insisting on blinding.

A more complex quality scoring system has been developed for systematic reviews by the Cochrane Collaboration Back Review Group. This has been adopted for use in assessing the quality of primary studies in hypnosis (work in progress, not reported here) and is included here as a quick reference for the reader who is interested in assessing the quality of individual studies. For advice regarding the application of these criteria, see van Tulder et al.

Quality assessment as utilized by the Cochrane Collaboration Back Review Group is as follows:

- **Patient selection**
  a) Were the eligibility criteria specified?
  b1) Was a method of randomization performed?
  b2) Was the treatment allocation concealed?
  c) Were the groups similar at baseline regarding the most important prognostic indicators?

- **Interventions**
  d) Were the index and control interventions explicitly described?
  e) Was the care provider blinded to the intervention?
  f) Were co-interventions avoided or comparable?
  g) Was the compliance acceptable in all groups?
  h) Was the patient blinded to the intervention?

- **Outcome assessment**
  i) Was the outcome assessor blinded to the intervention?
Levels of evidence and empirically supported treatments

Although the quality of individual studies is important, the value of other characteristics of evidence such as level, relevance and strength has also been formally recognized. Attempts to discriminate between types of evidence have resulted in ‘league tables’, as illustrated in Table 1.

Some confusion is possible because of the similarity of alternative rating systems. For example, the system shown in Table 2 uses criteria and terminology that have similarities, but also some differences.

In Australia, the National Health and Medical Research Council has adapted the US descriptors (Table 2) to give a further version (Table 3). The National Health and Medical Research Council guidelines about using the best possible evidence moved away from a reliance on the opinions of experts towards reliance on the best scientific evidence. ‘Best available’ has come to be understood in terms of the level, quality, relevance and strength of evidence. Each of these parameters will be considered in the present attempt to describe the best available evidence relating to hypnotic analgesia. Level refers to the study design used in attempts to minimize bias (as illustrated by Tables 1–3). Quality refers to attempts to minimize bias within a study (see the earlier discussion of RCTs as a gold standard). Relevance refers to the extent to which the findings from a study could be applied in other clinical settings to different patients. Strength refers to the magnitude of effect sizes or to other statistical measures, such as statistical and clinical effectiveness, and the width of confidence levels.

The evidence-based movement in counselling and psychology

The term ‘empirically supported treatment’ (EST), which is similar in meaning to the concept of ‘evidence-based medicine’ (previously defined), has become prominent in psychology and counselling.

Chambless and Hollon have described a scheme for determining whether a particular treatment may be considered as established in efficacy or to be possibly efficacious. According to these guidelines, for a treatment to be considered efficacious, there must be at least two between-group design experiments, conducted by at least two different investigative teams, showing

<table>
<thead>
<tr>
<th>Type</th>
<th>Evidence source</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Strong evidence from at least one systematic review of multiple well-designed RCTs</td>
</tr>
<tr>
<td>II</td>
<td>Strong evidence from at least one properly designed RCT of appropriate size</td>
</tr>
<tr>
<td>III</td>
<td>Evidence from well-designed trials without randomization, single group pre-post, cohort, time series, or matched case-controlled studies</td>
</tr>
<tr>
<td>IV</td>
<td>Evidence from well-designed nonexperimental studies from more than one centre or research group</td>
</tr>
<tr>
<td>V</td>
<td>Opinions of respected authorities, based on clinical evidence, descriptive studies or reports of expert committees</td>
</tr>
</tbody>
</table>
superiority to a no-treatment control condition, an alternative treatment, or a placebo, or indicating equivalence to a previously established efficacious treatment. More details on the criteria for empirically supported psychological treatments are given below:

1) Comparison with a no-treatment control group, alternative treatment group, or placebo: (a) in a randomized controlled trial, controlled single case experiment, or equivalent time-samples design; and (b) in which the EST is statistically significantly superior to no treatment, placebo, or alternative treatments or in which the EST is equivalent to a treatment already established in efficacy, and power is sufficient to detect moderate differences.

2) These studies must have been conducted with: (a) a treatment manual or its logical equivalent; (b) a population, treated for specified problems, for whom inclusion criteria have been delineated in a reliable, valid manner; (c) reliable and valid outcome assessment measures, at minimum tapping the problems targeted for change; and (d) appropriate data analysis.

3) For a designation of ‘efficacious’, the superiority of the EST must have been shown in at least two independent research settings (sample size of three or more at each site in the case of single case experiments). If there is conflicting evidence, the preponderance of the well-controlled data must support the EST’s efficacy.

4) For a designation of ‘possibly efficacious’, one study (sample size of three or more in the case of single case experiments) suffices in the absence of conflicting evidence.

5) For a designation of ‘efficacious and specific’, the EST must have been shown to be statis-
Assessment of the quality of integrative literature reviews

A well as original research reports, the literature includes narrative reviews, systematic reviews and meta-analyses, in which reviewers have attempted to synthesize other published works and derive an overall conclusion about treatment outcome. Reviews help readers to cope with the exponentially increasing number of primary publications. They attempt to provide synthesis and integration to a field and supply the reader with a summary picture. Understandably, they tend to be relied on by readers to provide an authoritative overview. This faith is illustrated in the following quotation: ‘systematic reviews and large randomised trials constitute the most reliable sources of evidence we can muster. Put simply, they are the best chance we have to determine what is true’ (p. 2).

McQuay and Moore have warned that systematic reviews of inadequate quality may be worse than none because faulty decisions may be made with unjustified confidence. Quality control in the systematic review process is vital. The quality of reviews varies, however, in much the same way as the quality of original research articles varies. McQuay and Moore noted that, in order for reviews to be valid, they have to be systematic and ‘to be systematic, qualitative or quantitative, they need to include all relevant randomised controlled trials (RCTs). Identifying all the relevant trials is a fundamental challenge which is easily underestimated’ (p. 7).

Comprehensive searches are typically compromised owing to time and cost factors. Consideration of these issues led McQuay and Moore to describe the process whereby a citation database was developed. There were three phases:

1) Definition of inclusion criteria (this may, for example, restrict consideration to publications of RCTs);
2) Identification of reports (this may include a clear description of the database search strategies used, together with information about any manual journal searches, requests to authors for information, etc.);
3) Information management (this may include a description of how the citation database is managed; e.g. software packages such as End Note, Pro-Cite, Reference Manager).

In order for a reader to be able to judge the merits of a systematic review, it is important that the phases identified above are reported in sufficient detail in the review itself. This step is often overlooked and many published reviews fail to provide a thorough description of their search strategies.

Once all relevant reports have been assembled, McQuay and Moore argued that the next step is an assessment of their quality (see also Jadad et al.). This is important because they found that lower-quality reviews tended to reach more favourable conclusions about treatment outcome. For example:

‘in a systematic review of Transcutaneous Electrical Nerve Stimulation (TENS) in postoperative pain, 17 reports on 786 patients could be regarded unequivocally as RCTs in acute postoperative pain. Fifteen of these RCTs demonstrated no benefit of TENS over placebo. Nineteen reports had pain outcomes but were not RCTs, in 17 of these 19, TENS was considered by their authors to have had a positive analgesic effect’ (p. 32).

We have already considered ways of assessing primary studies through the use of a quality scale. It is important, too, that review studies are assessed for their quality.

In their own systematic review of pain interventions, McQuay and Moore used a quality assessment system devised and validated by Oxman and Guyatt (Table 4). The Oxman and Guyatt index of scientific quality will be used in this report to assess the quality of published reviews of hypnosis and pain (each noted as ‘[Review]’ in the reference list). A newly developed methodological technique (named here as ‘a systematic meta-review’) will be used with the following goals:
Method

Reviews of hypnosis and pain

Identification of reports


As part of broader study of the hypnosis and pain literature, another resultant set of over 1200 references was also hand checked for review studies. A prominent hypnosis discussion list was used to appeal for additional information about hypnosis and pain (HYPNOSIS-request@MAELSTROM.STJOHNS.EDU). Hard copies of all potentially eligible reports were obtained. In addition, journal reference lists were scrutinized for further studies and the author’s own files were also utilized.

Information management

Search data were downloaded and managed with End Note software.

Inclusion criteria

This study focuses on reviews of hypnosis and clinical pain (not primary studies) published after Turner and Chapman’s sceptical 1982 review. Studies of experimentally induced pain were not included. Systematic reviews, narrative reviews and meta-analyses were included. Following the Cochrane Collaboration Back Review Group guidelines for systematic reviews, reviews based on nonrandomized controlled clinical trials were included in the present review, owing to insufficient available evidence from RCTs alone.

The Malone and Strube meta-analysis of non-medical treatments for chronic pain was not included because, although it did include 16
hypnosis studies, only one provided sufficient data to calculate an effect size. Even though this was a large effect size (2.67), no general conclusions should be drawn from just one study. A widely quoted review study (which concluded that ‘hypnosis appears to be of unique value in the treatment of clinical pain’ (p. 215)) was also not included because it was published in the same year as the critical review by Turner and Chapman and was not based on well-controlled studies of pain.

Some reviews that considered only case studies were excluded (e.g. Maline and Strube). Conceptual reviews that mentioned hypnosis but in which the authors did not explicitly attempt to evaluate efficacy and pain reviews that included only a brief description of hypnotic studies were also excluded. Conference articles and dissertations were not included, neither were non-English language publications nor reviews that were highly derivative of earlier reviews (Spink was based on a well-known hypnosis and pain text, Hilgard and Hilgard; and Weir was based on Van der Does and Van Dyck).

**Quality assessment**

Articles were subjected to quality scoring according to the Oxman and Guyatt index (Table 5). Following the procedure adopted by McQuay and Moore, each study was evaluated twice and a consensus score was obtained. A difficulty with this index is that the last item (see Table 4), requires a subjective judgement to be made about overall quality, with little guidance about how best to reach the decision. Consequently, in the present study, the last item was not used and the quality score became a mark out of nine. Predictably, narrative reviews usually achieved very low quality scores (compared with systematic reviews) because the narrative review process does not necessarily highlight methodological rigour in the selection or description of studies. It should be noted that quality in the present context refers to scientific quality and does not measure literary quality, importance, relevance, originality or other attributes of reviews.

**Results**

Quality scores ranged from 0 to 8 on a 9-point scale modified from the Oxman and Guyatt index of scientific quality. The median score was 1 (mean = 2.8, mode = 1). There was no apparent association between year of publication and quality score ($r = 0.39, p = 0.07$). A sensitivity analysis showed that there was no consistent relationship between the quality of reviews and the type of conclusion. The mean quality score for reviews rated as having negative or neutral conclusions was similar to the mean score for reviews rated as having positive conclusions (studies selected for closer analysis, as given in Tables 5 and 6; study nos 1–5, 8–10, 13, 14, 18, 22, 23). For example, a review by Ellis and Spanos, which received a quality score of only 1, was careful to conclude that, owing to methodological weaknesses, it was impossible to draw any conclusions about efficacy.

Included review studies were categorized according to the type of pain considered. Table 6 presents the quality scores and conclusions for each review. Studies that assessed pain relating to cancer or the invasive medical procedures associated with the treatment of cancer, especially in children, were the most frequently reported. Given the high incidence of chronic pain and the development of multidisciplinary clinics for its treatment, there were surprisingly few reviews of hypnosis and chronic pain.

**Description of the included reviews**

**Various types of pain**

Montgomery et al. produced the only meta-analysis of the effects of hypnosis on pain found in the literature. This classified hypnosis as a ‘well established treatment’. They identified relevant studies from existing reviews of the literature and from a PsycLIT data base search for the years 1974 to 1997. They restricted inclusion to studies that used hypnosis to attempt to reduce pain, studies that included a no-treatment or standard-treatment control group, and studies that included sufficient data to allow the calculation of effect sizes. This resulted in the inclusion of 18 studies and the calculation of 27 effect sizes. Overall the results indicated a moderate to large effect size ($d = 0.74$). The authors concluded that
### Table 5  Methodological features of the included studies (modified Oxman and Guyatt index of scientific quality\textsuperscript{14})

<table>
<thead>
<tr>
<th>Feature</th>
<th>Study no.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23</td>
</tr>
<tr>
<td>Search methods stated</td>
<td>1 0 0 1 1 1 0 0 0 0 0 0 1 0 0 0 0 0 1 1 0 0 0</td>
</tr>
<tr>
<td>Comprehensive search</td>
<td>1 0 0 1 1 1 0 0 0 0 0 0 1 0 0 0 0 0 1 1 0 0 0</td>
</tr>
<tr>
<td>Inclusion criteria reported</td>
<td>1 0 0 1 1 1 1 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 0</td>
</tr>
<tr>
<td>Selection bias avoided</td>
<td>1 0 0 1 1 1 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0</td>
</tr>
<tr>
<td>Validity criteria reported</td>
<td>0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0</td>
</tr>
<tr>
<td>Validity assessed</td>
<td>0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0</td>
</tr>
<tr>
<td>Combination methods reported</td>
<td>1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0</td>
</tr>
<tr>
<td>Findings combined appropriately</td>
<td>1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0</td>
</tr>
<tr>
<td>Conclusions supported by data</td>
<td>1 0 0 1 1 1 1 1 1 1 1 1 1 1 0 1 1 0 1 1 1 1 1</td>
</tr>
<tr>
<td>Total (max = 9)</td>
<td>7 0 0 7 7 7 4 1 1 1 1 1 7 1 1 0 3 2 8 3 1 1 1</td>
</tr>
</tbody>
</table>

1 = design feature was present; 0 = design feature was absent
Table 6  Review study quality scores and conclusions, using the modified Oxman and Guyatt index

<table>
<thead>
<tr>
<th>Study no. and reference</th>
<th>Quality score (/9)</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Various types of pain (3 reviews):</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1) Montgomery <em>et al</em>. 2000</td>
<td>7</td>
<td>'meta-analysis of 18 studies revealed a moderate to large hypnotanalgesic effect, supporting the efficacy of hypnosis for pain management' (p. 138)</td>
</tr>
<tr>
<td>2) Holroyd 1996</td>
<td>0</td>
<td>'controlled studies comparing hypnosis to other psychological interventions for pain frequently have shown hypnosis to be equally or more effective' (p. 34)</td>
</tr>
<tr>
<td>3) National Institute of Health Assessment Panel 1996</td>
<td>0</td>
<td>There was 'strong evidence for the use of hypnosis in alleviating the pain associated with cancer' (p. 313). There were other data 'suggesting the effectiveness of hypnosis in other chronic pain conditions, which include irritable bowel syndrome, oral mucositis, temporomandibular disorders, and tension headaches' (p. 315)</td>
</tr>
<tr>
<td><strong>Cancer/invasive medical procedures (9 reviews):</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4) Sellick and Zaza 1998</td>
<td>7</td>
<td>'there is much support for the specific use of hypnosis in managing pain associated with medical procedures and some support for its use in managing chronic cancer pain' (p. 13)</td>
</tr>
<tr>
<td>5) Pan <em>et al</em>. 2000</td>
<td>7</td>
<td>'support group therapy can improve pain in advanced cancer patients. Hypnosis can enhance such pain relief. Evidence strength: Level 2' (p. 385)</td>
</tr>
<tr>
<td>6) Milling and Costantino 2000</td>
<td>7</td>
<td>'it is perhaps a bit too early in the development of our research knowledge base to be able to say that any child hypnosis intervention has attained the milestone of meeting EST criteria for an efficacious therapy. However, based on encouraging preliminary evidence it seems that clinical hypnosis with children has taken its first steps towards empirical support' (p. 133)</td>
</tr>
<tr>
<td>7) Trijsburg <em>et al</em>. 1992</td>
<td>4</td>
<td>'positive effects were found with respect to specific symptoms such as anxiety, pain, nausea and vomiting' (p. 508)</td>
</tr>
<tr>
<td>8) Genuis 1995</td>
<td>1</td>
<td>'the consistency of these findings indicates the usefulness of hypnosis as an effective intervention for helping children and adolescents control anxiety and pain, and emesis associated with both the illness of cancer and certain medical procedures used in treating cancer patients' (p. 322)</td>
</tr>
<tr>
<td>9) Ellis and Spanos 1994</td>
<td>1</td>
<td>'Due to the imprecision with which the construct of hypnosis was operationalized and the lack of standardization across studies, however, it is impossible to draw definitive conclusions' (p. 103)</td>
</tr>
<tr>
<td>10) Liossi 2000</td>
<td>1</td>
<td>'clearly hypnosis has been shown in a number of studies to reduce the distress of children with cancer undergoing a variety of stressful procedures and chemotherapy' (p. 130)</td>
</tr>
<tr>
<td>11) Spiegel and Moore 1997</td>
<td>1</td>
<td>'whatever the mechanism, hypnotic analgesia is efficacious' (p. 1182)</td>
</tr>
<tr>
<td>12) Sutters and Miaskowski 1992</td>
<td>1</td>
<td>'psychological methods appear to be beneficial in reducing children’s pain and anxiety during invasive procedures such as BMA [bone marrow aspiration] and LP [lumbar puncture]' (p. 467)</td>
</tr>
<tr>
<td><strong>Chronic pain (4 reviews):</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13) Kirsch <em>et al</em>. 1995</td>
<td>7</td>
<td>'hypnosis can be a useful adjunct to cognitive behavior therapy for a wide range of problems' (p. 219)</td>
</tr>
<tr>
<td>14) Large 1994</td>
<td>1</td>
<td>'the few controlled studies published evaluating the efficacy of hypnosis for chronic pain suggest that it is an effective therapy in reducing subjective pain experience, but no more so than relaxation or autogenic training' (p. 237)</td>
</tr>
<tr>
<td>15) Spinhoven 1987</td>
<td>1</td>
<td>'in the only controlled study of this subject, the effectiveness of hypnosis and of relaxation training appeared comparable' (p. 125)</td>
</tr>
</tbody>
</table>
Hypnosis and clinical pain

Hypnotic suggestion is an effective analgesic, based on the combined analyses from more than 900 participants. They found that, for 75% of the sample population, hypnosis provided substantial pain relief and there were no differences in effectiveness between clinical patients and healthy volunteer samples (experimental pain) and that the effects of hypnosis were mediated in the expected direction by measured hypnotizability.

To make the case that hypnosis was efficacious according to the criteria described by Chambless and Hollon, Montgomery et al. used as examples the studies with cancer patients and burn patients. The additional requirement that studies should utilize treatment manuals was not met, but Montgomery et al. argued that, since the hypnotic treatment procedures are ‘readily available’ in the literature and the study procedures are described within the research articles, their conclusion is nonetheless acceptable, and that ‘meta-analysis of 18 studies revealed a moderate to large hypoanalgesic effect, supporting the efficacy of hypnotic techniques for pain management’ (p. 138).

Holroyd noted that eight controlled studies comparing hypnosis to other psychological interventions for pain have shown hypnosis to be equally or more effective, but provided no assessment of the methodological rigour of the contributory studies. Her review then focused on three studies claimed to ‘exemplify’ the effectiveness of hypnosis relative to other therapies. No further rationale for the selection of these three studies was provided.

Holroyd first drew attention to a study of fibromyalgia, which showed that hypnosis reduced pain and other symptoms in a fibromyalgia group more than in a comparison group using...
massage plus relaxation. The second study, which used random allocation to groups, showed that hypnosis was superior to no treatment or placebo in relieving pain from the treatment of burns. The hypnosis group demonstrated pain to 54% of baseline compared with 84% and 86% of baseline for the placebo and no-treatment groups respectively. The third study concerned hypnosis for pain in bone marrow transplantation. Hypnosis was compared with cognitive behavioural treatment and placebo. The hypnosis patients had significantly less pain (and other symptoms) than the patients in the other two groups, while using less pain medication.

The National Institute of Health Assessment Panel on the Integration of Behavioral and Relaxation Approaches is a 12-member non-advocate multidisciplinary panel (USA). They considered literature search data from 23 experts at the National Institute of Health Technology Assessment Conference with the objective of providing physicians with a responsible assessment of behavioural and relaxation approaches to the treatment of chronic pain and insomnia. This group concluded that ‘there is strong evidence for the use of hypnosis in alleviating the pain associated with cancer’ (p. 313) and that there were other data ‘suggesting the effectiveness of hypnosis in other chronic pain conditions, which include irritable bowel syndrome, oral mucositis, temporomandibular disorders, and tension headaches’ (p. 315).

Cancer/invasive medical procedures

Sellick and Zaza provided a review of five nonpharmacological strategies for managing cancer pain. These included hypnosis (Table 7 presents a modified version of these data). This review is distinctive in that it used a quality assessment tool to assist the evaluation of contributory studies. The authors selected only RCTs from an extensive literature search of Medline, CINAHL and PsycINFO. The hypnosis search returned 12185 references, including 37 RCTs of hypnosis and chronic pain and six RCTs of hypnosis and cancer pain. They concluded that:

although the effectiveness of hypnosis in the management of cancer pain has not been examined extensively, there is much support for the specific use of hypnosis in managing pain associated with medical procedures and some support for its use in managing chronic cancer pain (p. 13).

Pan et al. reviewed complementary and alternative medicine in patients at the end of life, including effects on pain. Their study focused on 21 studies after initially considering 619 citations. Of these, only 14 studies addressed pain and only a single study was reported that focused on the effects of hypnosis. This study was an unblinded RCT of 58 women with advanced breast cancer. Group therapy helped these women in terms of pain relief compared with a standard treatment control and a third, hypnosis, group achieved pain relief significantly greater than in even the successful group therapy women. On the basis of this single study the authors concluded that hypnosis can relieve pain in advanced cancer patients and they rated the strength of this evidence at level 2 (United States Preventive Services task Force criteria; i.e. ‘evidence obtained from at least one properly designed randomised controlled trial’). A strength of this review was that it considered a hierarchy of evidence using a best evidence approach; data were initially sought from meta-analyses and RCTs; then, if such data were not available, other controlled trials and then finally case series were considered.

Milling and Costantino reviewed hypnosis for children across a wide range of applications, including reduction of acute pain. They considered five studies involving acute pain in children. When these studies were considered against the Chambless and Hollon criteria for establishing an EST, the authors noted that, although random assignment to condition was one of the strengths of these reports, no study utilized a treatment manual as required by the guidelines (a manual is useful to operationalize the intervention clearly, and to allow replication and comparison across studies). This failure was described as a serious obstacle to obtaining EST status. These authors noted that the literature on child hypnosis is predominantly composed of anecdotal case histories and uncontrolled research studies, but they considered the field to show much promise (especially the areas relating to acute pain and distress) and they concluded that:

it is perhaps a bit too early in the development of...
our research knowledge base to be able to say that any child hypnosis intervention has attained the milestone of meeting EST criteria for an *efficacious* therapy. However, based on encouraging preliminary evidence it seems that clinical hypnosis with children has taken its first steps towards empirical support (p. 133).38

Trijsburg *et al.*39 reviewed psychological treatments of cancer patients. They considered 22 studies with at least one control group, three of which utilized hypnosis.72,77,80 In their assessment of the hypnosis studies, these authors concluded that ‘positive effects were found with respect to specific symptoms such as anxiety, pain, nausea and vomiting’ (p. 508).

Genuis40 reviewed nine studies concerning applications of hypnosis for cancer, including the relief of pain in paediatric patients. In seven of these studies pain scores were assessed; there was a significant reduction in the pain experienced by patients after hypnosis.77,82,83,87–90 Genuis concluded that the consistency of these findings indicates the usefulness of hypnosis. He was critical of the methodological shortcomings of the included studies but failed to provide sufficient information about the methodology of contributing studies (e.g. it was not always clear which studies used random allocation to groups). Four of the included studies compared hypnosis with other psychological treatment methods; two of these showed hypnosis to be of superior value.77,91 Genuis suggested that hypnotizability data should be included in comparative reports, presumably in case it mediates positive outcome that is independent of the actual use of formal hypnosis. He suggested that controlling for attention given across groups is insufficient and that a measure of the rapport developed with patients is an important control that should be used in future studies. Implicit in this recommendation is the idea that positive effects may be an artefact of good rapport even when the attention given has been balanced across groups.

Ellis and Spanos33 reviewed studies that used a variety of cognitive behavioural interventions in children facing bone marrow aspiration and lumbar puncture. They identified six studies that utilized hypnosis.77,82,83,86,89,90 They noted the difficulty in comparing studies that used hypnotherapy because of the lack of standardized procedures in its clinical use and because the definition of a situation as hypnotic is not always comparable across studies. For example, they commented that at times an hypnotic induction followed by breathing exercise, distraction, imagery, suggestion or therapist support were all labelled as hypnosis. After reviewing the studies, which generally reported reduced pain after hypnosis, they noted the difficulty in attributing the changes to the hypnosis and not to some other aspect of the situation. To illustrate the argument, they noted that in Kuttner *et al.*’s study86:

the hypnosis group used imagery, suggestion and therapist support while the control group used bubble blowing, counting, puppet play and pop-up books. To compare precisely the effects of hypnotic procedures on distress it would have been more useful to have both groups engage in the same activities, e.g., bubble blowing, pop-up books, guided imagery, etc, with these activities preceded in one group by the hypnotic induction. Any differences between the groups could then have been unambiguously attributed to the one variable that differed between them, the hypnotic induction ritual (p. 102).33

Their conclusion was that the studies seemed to indicate that hypnosis offered no additional benefits to those achievable with distraction, imagery and other strategies, but they also acknowledged that definitive conclusions were impossible to reach because of the imprecision with which hypnosis was operationalized and because of a lack of standardization across studies. They called for improved methodological quality in studies before a conclusion is drawn.

Liossi41 reviewed applications of clinical hypnosis in paediatric oncology, including attempts at pain control. She reviewed eight studies using hypnosis (the same studies were also included in a broader review of cancer pain92). In all of the studies considered, hypnosis was associated with reduced pain. When control groups were used, hypnotic effects were typically equivalent to alternative interventions, which included behavioural techniques,77 play therapy,82 distraction,86 cognitive strategies83 and cognitive behavioural training93. The three studies that did not use control groups were Hilgard and LeBaron89 (baseline post-test), Kellerman *et al.*’90 (baseline post-test), and Hawkins *et al.*94 (parallel group).
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Sample</th>
<th>Outcome measure</th>
<th>Results</th>
</tr>
</thead>
</table>
| Spiegel and Bloom 1983<sup>80</sup> | 3 groups:  
G1: 90 min weekly support group with 5–10 min hypnosis session  
G2: 90 min weekly support group without hypnosis  
G3: Usual treatment | 54 women with metastatic breast cancer; mean age 55.5 yr | 10-point scales of pain sensation, frequency and duration, and of suffering; POMS and medication intake | Significant difference in mean pain sensation in G1 (0.008) and G2 (0.05) compared with G3 (0.77) |
| Syrjala et al. 1992<sup>66</sup> | 4 groups:  
G1: Usual treatment  
G2: Therapist contact  
G3: CBT (excluding imagery)  
G4: Hypnosis | 45 patients undergoing preprocessing | BSI, SIP, oral pain VAS, nausea VAS, opioid use | G4 had significant decrease in oral pain |
| Syrjala et al. 1995<sup>81</sup> | 4 groups:  
G1: Usual treatment  
G2: Therapist support  
G3: Relaxation and imagery  
G4: CBT (including imagery) Groups stratified for sex | 94 patients undergoing preprocessing | Oral pain VAS, nausea VAS, opioid intake, oral mucostis severity, SCL-90R, post-treatment evaluation | G3 and G4 had significant reduction in pain; no added benefit for G4 over G3 |
| Zelzer and LeBaron 1982<sup>77</sup> | 2 groups:  
G1: Hypnosis  
G2: Nonhypnotic behavioural techniques (nonimagery distraction) | 33 paediatric patients undergoing BMA or LP; age range 6–17 yr | 5-point scale of pain and anxiety | BMA patients: Pain significantly reduced in both groups; G1 effect size larger (decrease 1.5 out of 5) than that of G2 (decrease 0.66); anxiety significantly reduced in G1 only |
### Hypnosis and clinical pain

<table>
<thead>
<tr>
<th>Study</th>
<th>Groups</th>
<th>Participants</th>
<th>Measures</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Katz et al. 1987&lt;sup&gt;52&lt;/sup&gt;</td>
<td>2 groups:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: Hypnosis patients</td>
<td>36 paediatric</td>
<td>PBRS-&lt;i&gt;r&lt;/i&gt; scale, rating of anxiety (nurse-related), pain (patient-rated) and fear (patient-rated)</td>
<td>Significant differences in both groups for pain and fear</td>
<td></td>
</tr>
<tr>
<td>G2: Play patients</td>
<td>undergoing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wall and Womack 1989&lt;sup&gt;83&lt;/sup&gt;</td>
<td>2 groups:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1: Hypnosis (relaxation, visual imagery) patients</td>
<td>20 paediatric</td>
<td>Pain VAS, McGill, Pain Questionnaire, pulse and temperature readings</td>
<td>Pain significantly reduced in both groups</td>
<td></td>
</tr>
<tr>
<td>G2: Active cognitive strategy (i.e. distraction) undergoing</td>
<td>BMA or LP</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

POMS, Profile of Mood States; CBT, cognitive-behavioural therapy; BMT, bone marrow transplantation; BSI, Brief Symptom Inventory; SIP, Sickness Impact Profile; VAS, visual analogue scale; SCL-90R, Symptom Check List – 90 Revised; BMA, bone marrow aspiration; LP, lumbar puncture; PBRS-<i>r</i>, Procedural Behaviour Rating Scale
Liossi concluded that ‘clearly hypnosis has been shown in a number of studies to reduce the distress of children with cancer undergoing a variety of stressful procedures and chemotherapy’ (p. 130).41

Spiegel and Moore,42 in a review of imagery and hypnosis for the treatment of cancer, asserted that hypnotic analgesia ‘is efficacious’. They described hypnosis as superior to a ‘control condition of sympathetic attention alone in children undergoing painful procedures’, ‘producing a 50% pain reduction among patients with metastatic breast cancer’ in a randomized prospective study, and being ‘more potent than either placebo analgesia or acupuncture analgesia’ (p. 1182). These conclusions were based on the works of Zeltzer and LeBaron,77 Spiegel and Bloom,80 McGlashen et al.95 and Knox and Shum96 respectively.

Sutters and Miaskowski43 reviewed psychological strategies for the management of pain in children with cancer. They noted that these have focused almost exclusively on acute pain associated with painful medical procedures rather than on chronic pain attributable to the disease process. Procedures receiving most attention were bone marrow aspiration and lumbar puncture. They reviewed both uncontrolled studies and what they called ‘structured designs’ (these included RCTs) and found a lack of consistency in the results when hypnotic procedures were considered. Their conclusion was a moderate one saying that psychological methods ‘appear to be beneficial in reducing children’s pain and anxiety during invasive procedures’ (p. 467). Of the nine studies presented in their review, six involved hypnosis. A summary of these six studies is reproduced in Table 8.

Chronic pain

Kirsch et al.44 conducted a meta-analysis of 18 studies in which cognitive behavioural therapy was compared with the same therapy supplemented by hypnosis. They concluded that the average client receiving hypnosis was better off at the end of treatment than 75% of the clients receiving nonhypnotic treatment. Only two of the included studies focused on pain.55,99 These studies had effect sizes of –0.20 and 0.16 respectively, an inconsistent finding from which no firm conclusion should be drawn.

Large45 commenced his review with the caution that most published accounts of the use of hypnosis for pain are anecdotal and/or uncontrolled. He reviewed 12 studies,66,74,79,100–108 although it seems noteworthy that, of these 12, three came from one particular research group. His conclusion, based on ‘the few published controlled studies evaluating the efficacy of hypnosis for chronic pain’ was relatively restrained in suggesting that ‘the overall impression from these studies is that hypnosis is an effective therapy in the management of chronic pain’. No information is provided about the search strategies used to identify the included studies and there is no information to confirm that the literature review was exhaustive.

Spinhoven, in 1987,46 did not attempt to review outcome studies per se; instead, he was interested in discussing the factors relating to positive outcomes, in particular whether organic versus psychogenic pains were differentially responsive and what were the relationships between psychosocial problems and low back pain. He noted that only one controlled study of hypnosis for low back pain had been reported at that time,109 which compared randomly allocated hypnosis and relaxation outpatient groups and found that, although both approaches were effective for improving pain, depression and medication use, and that both were more effective than placebo, there were no differences between the two. As well as being the only available RCT on hypnosis and chronic low back pain, this study is distinctive in that it utilized treatment manuals to standardize the treatments used (as has been advocated recently by Chambless and Hollon13). The authors noted a cost advantage compared with inpatient treatment (a comparison of $7500 versus $700) but unfortunately did not assess hypnotizability, so the relative contributions of hypnotic processes and relaxation cannot be determined.

Godoy and Araoz47 claimed that between 1989 and 1997 there were 672 studies of hypnotic outcomes in psychiatry reported in journals (the search strategy is not described), of which 14 related to chronic pain, 11 of these to migraine. The 14 studies were not identified, there was no significant analysis, and the conclusion that ‘the
### Table 8  Studies investigating psychological interventional strategies for the management of pain in children with cancer

<table>
<thead>
<tr>
<th>Reference</th>
<th>Purpose</th>
<th>Sample and design</th>
<th>Major findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hilgard and LeBaron 1982&lt;sup&gt;89&lt;/sup&gt;</td>
<td>To examine the use of hypnosis for relief of acute pain and anxiety relating to BMA</td>
<td>24 children with cancer undergoing BMA, aged 6–9 yr</td>
<td>15 of the 24 patients were successful in reducing self-reported pain below baseline level with hypnosis. Hypnotic ability was found to be an important factor in successful reduction of self-reported pain(&lt;i&gt;p&lt;/i&gt; &lt; 0.05)</td>
</tr>
<tr>
<td>Hockenberry 1988&lt;sup&gt;97&lt;/sup&gt;</td>
<td>To demonstrate the use of various relaxation techniques by paediatric patients experiencing discomfort and stress from cancer treatment</td>
<td>15 children and adolescents with cancer, aged 6–15 yr</td>
<td>Findings suggest that relaxation techniques (e.g. distraction, imagery, muscle relaxation, hypnosis) can offer effective coping strategies for children with cancer who are undergoing painful, invasive procedures</td>
</tr>
<tr>
<td>Olness 1981&lt;sup&gt;87&lt;/sup&gt;</td>
<td>To describe the clinical experience of patients using imagery as an adjunct therapy in the treatment of malignancy</td>
<td>25 children and adolescents with cancer, aged 3–18 yr; followed up for 2–5 yr (or until their death)</td>
<td>19 patients demonstrated symptom relief in association with imagery exercises, suggesting that this modality may be a valuable adjunct, especially for those who begin exercise at the time of diagnosis. Children aged 5–11 yr who used imagery demonstrated more rapid pain control than adolescents</td>
</tr>
<tr>
<td>Wall and Womack 1989&lt;sup&gt;83&lt;/sup&gt;</td>
<td>To compare the efficacy of standardized instruction in hypnosis or active cognitive strategy for providing relief from procedure-induced pain and anxiety</td>
<td>20 children and adolescents with cancer, aged 5–18 yr; grouped by age 5–11 yr (younger) and 12–18 yr (older), then randomly assigned to treatment conditions Experimental</td>
<td>Standardized instruction in hypnotic and active cognitive strategies (i.e. choice of distraction: concentration on motoric or on sequential information) produced significant reductions in reported pain. Neither technique provided for anxiety relief</td>
</tr>
<tr>
<td>Zeltzer and LeBaron 1982&lt;sup&gt;77&lt;/sup&gt;</td>
<td>To compare hypnosis to nonhypnotic behavioural techniques for efficacy in reducing BMA and LP for children with cancer</td>
<td>52 children and adolescents with cancer, aged 7–17 yr; randomized to hypnosis and nonhypnosis groups Experimental</td>
<td>Both types of behavioural intervention resulted in pain reduction for 21 patients receiving intervention during BMA, although the hypnosis group demonstrated greater pain reduction than the nonhypnosis group (&lt;i&gt;p&lt;/i&gt; &lt; 0.0001). For 17 patients receiving intervention for LP, both hypnosis and nonhypnotic strategies were equally effective in reducing pain (&lt;i&gt;p&lt;/i&gt; &lt; 0.01) and anxiety (&lt;i&gt;p&lt;/i&gt; &lt; 0.02)</td>
</tr>
<tr>
<td>Zeltzer and LeBaron 1982&lt;sup&gt;98&lt;/sup&gt;</td>
<td>To compare hypnosis and nonhypnotic behavioural techniques for efficacy in reducing pain and anxiety for children and adolescents undergoing BMA and LP</td>
<td>45 children and adolescents with cancer, aged 6–17 yr; 27 undergoing BMA and 22 LP; assigned randomly to hypnosis or nonhypnosis intervention groups, and matched for age and disease type Experimental</td>
<td>Findings suggested that hypnosis was more effective than nonhypnotic techniques for both pain and anxiety reduction. Variation within both groups on the extent to which they were helped by either technique suggests other potentially important intervening variables</td>
</tr>
</tbody>
</table>

BMA, bone marrow aspiration; LP, lumbar puncture
use of hypnosis has been relevant to the symptomatic improvement of these patients’ was not supported by any evaluation of the contributory studies.

**Burns**

Patterson *et al.* reviewed the effects of hypnosis on burn pain. They considered both anecdotal reports (single-case designs, multiple-case designs, and uncontrolled studies) and controlled studies. They considered six controlled studies, four of which were from Patterson’s own research group.

The first study considered, by Hammond *et al.*, involved experimentally-induced burn pain, which responded to hypnotic analgesia suggestion in that none of the participants reported pain in the thigh that was the focus of suggestion. In Wakeman and Kaplan’s study, which was described by Patterson *et al.* as the first well-controlled study of hypnotic analgesia in burn patients (Crasilneck *et al.* are usually cited as having provided the first description of the application of hypnosis to burns), patients receiving hypnosis requested significantly less medication than a control group who received attention from a psychologist.

In the first study by Patterson *et al.* they improved methodological quality by using a standard replicable hypnotic induction, which showed that a hypnosis group produced significantly decreased visual analogue pain scores subsequent to a hypnotic intervention after three earlier baseline measurement periods. A no-treatment (historical control) group showed no such comparable decreases of pain over time.

The next study from this research team utilized a randomization approach, which compared three groups (active hypnosis plus medication, placebo hypnosis plus medication, medication alone). They used only patients who reported severe pain (greater than 5 on a 0 to 10 visual analogue scale). Pain was initially recorded for two days for baseline levels. Patients in the active hypnosis group showed significantly decreased pain scores as determined by both patient and nurse ratings of pain.

The third study from Patterson and colleagues compared four randomly assigned groups (hypnosis alone, lorazepam alone, hypnosis and lorazepam, and opioid medication alone). All groups showed decreases in pain reports but no treatment group showed an effect superior to the others. It was postulated that low baseline pain might have made changes difficult to detect (a floor effect), and a subsequent study demonstrated that, when baseline pain was not considered, the experimental group (hypnosis) did not differ from the control group, but, when patients with high initial pain scores were considered separately, there was a significant difference between groups in favour of the hypnosis group.

In summarizing their review, Patterson *et al.* concluded that there was ‘compelling evidence’ for the analgesic effect that hypnosis can have on burn patients.

Van der Does and Van Dyck reviewed the use of hypnosis in the care of burn patients, including an emphasis on pain control. These authors noted the discrepancy between the dramatic results suggested by anecdotal reports and the more modest findings of systematic enquiries. They discussed the Wakeman and Kaplan study, which is notable for using random allocation to groups. In that study, medication levels were used as an indirect measure of pain and the experimental group utilized less medication. Van der Does and Van Dyck were critical of the methodology, however, noting that the experimental group might have been more willing to try to undergo a dressing change with less medication than usual because they were aware they had received a treatment intended to reduce pain. They concluded the review by suggesting that the use of hypnosis in alleviating burn pain is promising but ‘it remains to be demonstrated that this will be a specific rather than a general or placebo effect’ (p. 123).

**Gastrointestinal pain/irritable bowel syndrome**

Talley *et al.* reviewed psychological treatments, including hypnosis, for the treatment of irritable bowel syndrome. This study achieved the highest quality score (8/9). The authors utilized a comprehensive quality algorithm to assess the value of contributory studies. They searched Medline and PsycLIT for comparative studies and selected 14 that were appropriate, two of which were based on hypnosis. They subjected...
these studies to a very thorough quality evaluation based on: participant selection variables (population type, whether patient selection was consecutive, refusal and drop-out rates, and the definition of irritable bowel syndrome used); research protocol variables (randomization, baseline period, treatment method, type of control group, equivalence of contact time, trial length, blinding, and follow-up); measures (baseline measures, concurrent drug use, psychological assessments, compliance, outcome measures, and expectancy); and statistical adequacy (intention-to-treat analysis, adequacy of reporting, appropriateness of statistical analyses, study power, and clinical significance).

On an a priori basis, they selected eight of the many criteria described above as being the most critical for study validity and interpretability. These were: randomization, blinding of outcome assessors, irritable bowel syndrome definition, accounting for concurrent drug use, equivalence of contact time where appropriate, acceptability of outcome measures, use of control groups, and statistical adequacy. They determined that in order to qualify a study as methodologically acceptable it would have to exhibit six of eight quality criteria. The only report to achieve this standard was a hypnotherapy study,\textsuperscript{101} which showed that irritable bowel syndrome patients improved significantly more than controls. Talley \textit{et al.} were nonetheless critical of this study because it utilized nonconsecutive volunteers and largely excluded people who were over 50 years old. These restrictions affected the generalizability of the study results. Given that this study achieved the \textit{a priori} quality score, it is surprising that Talley \textit{et al.}\textsuperscript{49} concluded that ‘no trial to date has produced unequivocal evidence that psychological treatment is efficacious in irritable bowel syndrome’ (p. 285).

Compas \textit{et al.}\textsuperscript{50} reviewed psychological treatments in four areas of health psychology, which met the criteria for ESTs. One focus was on chronic pain and one of the treatment modalities for chronic pain was hypnosis (the article also covered operant behavioural treatment, cognitive behavioural treatment, biofeedback and psycho dynamic treatment). Curiously, the review selected only four pain conditions for analysis (rheumatic diseases, chronic pain syndromes (e.g. back pain), migraine headache, and irritable bowel syndrome). No data were reported for the use of hypnosis for rheumatic diseases, or chronic low back pain, or migraine headache. Hypnosis was discussed in terms of irritable bowel syndrome and the conclusion was that it was ‘possibly efficacious’; however, only two studies were cited. The first was a controlled study by Whorwell \textit{et al.},\textsuperscript{101} which showed that hypnotherapy significantly reduced abdominal pain and distension, and enhanced well-being. Patients undergoing hypnotherapy responded significantly better than a control group who received psychotherapy plus placebo medication. The second study\textsuperscript{115} showed that 20 of 33 patients improved with hypnotherapy, with 11 losing almost all of their symptoms. This study failed to employ a nonhypnosis control group. This review by Compas \textit{et al.}\textsuperscript{50} was disappointing in that it failed to review an adequate quantum of the published evidence before drawing conclusions. Only two empirical hypnosis studies were described despite the availability in the literature of many more.

Francis and Houghton\textsuperscript{51} provided an overview of the application of hypnotherapy in gastrointestinal disorders. They began with a series of rather dramatically successful case reports but also discussed two controlled trials from Whorwell and coworkers,\textsuperscript{101,105} both of which showed positive effects of hypnosis.

The therapeutic approach taken by the Whorwell group\textsuperscript{101,105,116} involved direct symptom control. Therapeutic suggestion focused on immediate physiological or symptomatic improvement rather than on any indirect attempts via distraction or anxiety reduction. Prior \textit{et al.}\textsuperscript{116} reported that hypnosis was able directly to affect rectal sensitivity as measured using the response to a balloon distension test. Francis and Houghton\textsuperscript{51} concluded that:

\textit{it appears that not only is hypnotherapy effective in the treatment of functional bowel disorders such as irritable bowel syndrome, but it may also play a useful role alongside conventional medication in the treatment of organic disorders, such as relapsing duodenal ulcer disease (p. 528).}

\textbf{Childbirth}

Baram\textsuperscript{52} considered applications of hypnosis in reproductive health care, including pain relief in labour and delivery. He described case studies,
including his own first ever experience with hypnosis in labour, which involved a distressed 15-year-old primigravida. She responded well to hypnotic induction (never before tried by the author with a patient) and ‘went through the remainder of her labour and delivery without pain, never asking for or requiring any medication, even for the episiotomy and repair’ (p. 40).52

After reviewing various controlled studies, Baram reported that there were only two well-designed randomized studies of hypnosis in labour and delivery. The first117 showed no advantage of hypnosis over psychoprophylaxis; the second118 initially trained women to control experimentally-induced ischaemic pain before applying this newly developed skill to labour. This study also assessed hypnotic susceptibility. Women who were highly susceptible and hypnotized had less post-partum depression, and hypnosis resulted in shorter stage 1 labour, less medication, more spontaneous deliveries, higher Apgar scores and less pain. Baram noted that ‘the best study to date concluded that hypnosis was a useful adjunct to traditional childbirth education’ (p. 41).52

**Headache**

Spinhoven53 reviewed hypnosis procedures for headache control. A comparison of hypnosis with other psychological treatments (biofeedback, relaxation training, autogenic training) showed that, although all treatments achieved some success in reducing headache, none of these procedures has consistently shown superior results compared with the others. The extent to which placebo responses are implicated cannot be determined in the absence of controlled studies comparing a credible placebo condition with hypnosis. In all of the reviewed studies in which hypnotizability was measured, there was a significant positive relationship between hypnotizability and therapeutic results (more than 350 patients), and Spinhoven concluded that ‘if we consider the level of hypnotizability rather than the details of the hypnotic procedure, it seems that headache patients who are highly hypnotizable benefit more from hypnosis in the reduction of headache’ (p. 190).53

**Discussion**

Chaves and Dworkin119 noted that, when the apparent success of acupuncture in controlling surgical pain was scrutinized carefully by study designs that included randomization, ‘sham’ acupuncture and patient blinding, the more rigorous designs were less likely to reveal a specific effect of acupuncture. They emphasized that this left unexplained the potency of psychosocial variables in attenuating pain. In the case of hypnosis, a similar effort to elicit specific effects of hypnosis is under way. The task is certainly unfinished because of methodological limitations, which have not always allowed specific effects to be easily distinguished from nonspecific effects such as social influence variables or placebo.

Although there are many references to poor methodological quality and the lack of RCTs in the hypnosis literature, hypnosis is not alone in this regard. A review of spinal cord stimulation for chronic low back pain120 noted that the lack of randomized trials precluded conclusions regarding the efficacy of spinal cord stimulation relative to other treatments, placebo or no treatment. Other examples can readily be found in almost any treatment field and the antidote to this clearly lies in the current emphasis developing in psychology and medicine on EST or EBM.

In the present review, well over 1000 articles relating to hypnosis and pain were identified from literature searches, yet the number of narrative and systematic reviews of this material was small and only one meta-analysis of the effects of hypnosis on pain was found. Consideration of this literature showed that not only does the quality of original research reports vary enormously but the quality of reviews of this research is just as variable.

**Best possible evidence**

In the light of the National Health and Medical Research Council guidelines12 about using the best possible evidence, some comments will be made about the level, quality, relevance and strength of evidence in the present review.

**Level (study design used in attempts to minimize bias)**

If an arbitrary cut-off quality score of 7/9 or above is used to assess reviews, it can be con-
cluded that level I evidence (evidence from a systematic review of all relevant trials) exists for the efficacy of hypnosis on pain in the form of the meta-analysis conducted by Montgomery et al. Level I evidence also exists for the effects of hypnosis on pain related to cancer/invasive medical procedure. There is a conspicuous lack of Level I evidence (at the quality level nominated) for the application of hypnosis in the treatment of such prevalent forms of pain as low back pain, headache, childbirth and burn pain.

**Quality (attempts to minimize bias within a study)**

Although there are some notable exceptions (e.g., Patterson’s studies of burn patients, the Whorwell group’s studies of IBS), poor methodological quality continues to dominate the hypnosis and pain literature. In particular, there remains a heavy reliance on case study reports or uncontrolled studies, and there is a notable lack of high-quality RCTs in specific areas (e.g., chronic pain, including low back pain). Although Milling and Costantino noted that random assignment to condition was a strength of the literature relating to hypnosis and cancer pain, the same cannot be said of the hypnosis and pain literature more generally. The lack of rigour in experimental design is a notable feature of the literature and is much complained about by reviewers. Some studies may have been well designed but they are not well reported; for example, it is not always clear from the publications whether random allocation to groups has occurred. Multiple other threats to internal validity exist, including the failure to include a credible placebo control condition.

**Relevance (the extent to which the findings from a study could be applied in other clinical settings to different patients)**

The pain of invasive medical procedures in children suffering from cancer has been studied often enough for generalization to be possible. Remarkably, the same cannot be said for various other types of pain, including chronic pains typically seen in multidisciplinary pain clinics.

The general failure to utilize treatment manuals means that replication is difficult and the extent to which generalization is possible is unknown. The use of treatment manuals is required to achieve EST status, so there is an urgent need to address this problem. Other threats to external validity include different definitions of hypnosis, and imprecision in the way that hypnosis is operationalized.

**Strength (the magnitude of effect sizes or other statistical measures such as statistical and clinical effectiveness and the width of confidence levels)**

Montgomery et al.’s meta-analysis showed a moderate to large effect size ($d = 0.74$) for the impact of hypnosis on pain. There is not yet in the hypnosis literature any tendency to report outcome data in a manner that will allow the combination of data across studies (e.g., L’Abbé plots, odds ratios, effect sizes, numbers needed to treat).

**Study goals**

The first goal was to produce a citation database of all available reviews, indicated in the reference list by ‘[Review]’. It is acknowledged that some reviews may not have been detected in spite of a comprehensive search effort. The database will be extended over time. The second goal was to assess the quality of available reviews and this information is given in Table 6.

Goal three was to test the hypothesis that poor-quality reviews would be more likely to produce positive conclusions, and to determine whether or not quality scores are useful to resolve conflicts between different systematic reviews. It is sobering, in this context, to note that previous research indicated that poor-quality reviews were more likely to reach conclusions favourable to the efficacy of the method under consideration than were good-quality reviews. Quality scores did not advance this issue because the relationship between review quality and type of conclusion noted by McQuay and Moore in TENS reviews did not apply in the present case. In general, with one or two exceptions, review authors were cautious in their conclusions and did not make extravagant claims for the efficacy of hypnosis.

The fourth goal was to determine whether hypnosis is efficacious in relieving clinical pain. McQuay and Moore warned against the simple conclusion that, if a majority of RCTs showed a treatment to be effective then it should be
accepted as such, on the basis that such simple vote counting may mislead as it ‘ignores the sample size of the constituent studies, the magnitude of the effect in the studies and the validity of their design even though they were randomised’ (p. 33). Montgomery et al.’s meta-analysis study was the one that best considered these variables. On the basis of an effect size of 0.74, that review concluded that hypnosis was efficacious for the treatment of pain.

The meta-analysis that demonstrated efficacy considered a range of pain types. When subcategories of pain are considered, the evidence is somewhat more sparse. There is some good-quality evidence for the efficacy of hypnotic treatment of pain relating to cancer or invasive medical procedures, and good studies have been conducted in the areas of burn pain and gastrointestinal-related pain. Surprisingly few good studies have been carried out in the areas of chronic pain, childbirth pain and headache pain.

**Promulgating lists of Empirically Supported Treatments**

A recent publication by Chambless and Ollendick has described efforts to increase the practice of evidence-based psychotherapies in a number of countries, with a particular focus on an American Psychological Association Task Force, which has published a series of reports and web-based material for the public (http://www.apa.org/divisions/div12/rev_est/index.shtml). Many specialist journals have also devoted special sections or editions to EST issues. Chambless and Ollendick themselves provided tables that list empirically validated psychological treatments for adults and for children.

Chambless and Ollendick described some caveats relating to the publication of EST lists. They acknowledged that nonappearance on a list may mean that: (1) the treatment in question has fared poorly in research trials; (2) the treatment has not been examined in research trials; or (3) the treatment was not reviewed. They further acknowledged that the fact that one ‘can identify and disseminate information about ESTs does not address arguments that it is ill advised to do so’ (p. 691). An example bearing on this potential for confusion arises with reference to hypnosis as a treatment for pain.

The Chambless and Ollendick table entries for the treatment of pain were informed by a review published by Wilson and Gill regarding the treatment of chronic pain, which specifically noted that psychological interventions for chronic pain have included hypnosis. The review then mentioned only two hypnosis studies specifically and therefore could not, with any authority, draw any conclusions about hypnotherapy. The review had aimed to assess the efficacy of broadly defined psychological interventions and with hypnosis specifically included, so the lack of a conclusion about hypnotherapy could be taken to mean that it had been found not to be an EST.

Chambless and Ollendick noted that additional criticisms of EST tables are based on concerns that the EST findings could be misused by managed care companies to disenfranchise practitioners of nonapproved psychotherapies, that such practitioners would be more vulnerable to malpractice suits, or that practice could become restricted to a limited number of treatments, thus precluding flexibility and clinical innovation.

An indication of the momentum behind the push for ESTs can be seen from statistics showing that the first American Psychological Association Task Force report on ESTs, in 1995, listed 25 treatments that met EST criteria. By 1998 the list included 71 treatments and, by 2001, 108 ESTs had been listed.

**Moving ahead**

The special edition of the *International Journal of Clinical and Experimental Hypnosis* published in April 2000 ended with a summary of the evidence relating to efficacy and made recommendations for the future. The summary authors noted that ‘hypnosis can now be considered a well-established treatment for pain’ (p. 242) and then made recommendations for the conduct of future studies, which included improving reporting considerations (define the population carefully, report the procedures in sufficient detail to permit replication, indicate clearly whether participants were randomly assigned to treatments, report hypnotic suggestibility of the samples, report complete descriptive data) and improving design considerations (ensure that the number of participants is adequate, conduct single- or multiple-case experiments, compare nonhypnotic
treatments with hypnotic inductions and suggestions added, conduct adequate follow-ups). To this, the present author adds the recommendation that authors of both primary studies and reviews could consider utilizing a quality scale such as one of those discussed here, to guide the development of their research effort.

The present review has been concerned only with assessing the evidence for the efficacy of hypnosis (whether it works) and the author has not commented on issues to do with how hypnosis may work. The parallel research effort into how hypnosis operates is also problematic in that no dominant theoretical model has yet emerged to explain the observations adequately. A number of frameworks for understanding hypnosis have been identified, ranging from the sceptical to the credulous. Broadly, these have been classified into two dichotomous views: the ‘special process’ (or ‘altered state’) perspective and the ‘social-psychological’ perspective. Many years ago, From and Shor described 12 key issues that an adequate theory of hypnosis would need to cover. That task, too, remains to be completed.

Conclusions

1) There is sufficient evidence, of sufficient quality, for a number of high-quality review studies to have concluded that hypnosis has demonstrable efficacy in the treatment of pain.

2) Although some high-quality reviews are supportive of the efficacy of hypnosis, review studies and the primary studies on which they are based have all been characterized by widespread variability in their quality. High-quality studies and high-quality reviews justify greater confidence in their findings, therefore it is hoped that dissemination of information about the measurement of quality, in both primary and review studies, will influence authors in their choice of study design.

References


99 McCauley JD, Thelen MH, Frank RG, Willard RR, Callen KE. Hypnosis compared to relaxation in the


