Inconsistencies in NICE guidance for acupuncture: reanalysis and discussion

Nicholas R Latimer, Aathish Chandrika Bhanu, David G T Whitehurst

ABSTRACT

Background Acupuncture received a positive recommendation in the National Institute for Health and Clinical Excellence (NICE) clinical guideline for low back pain (LBP). However, no such recommendation was forthcoming in the NICE clinical guideline for osteoarthritis (OA). Importantly, the two guidelines adopted different treatment comparators in their economic analyses of acupuncture; in the LBP guideline ‘usual care’ was used (with no consideration of placebo/sham interventions), whereas ‘sham acupuncture’ was the comparator in the OA guideline.

Objective To analyse the implications of using different control group comparators when estimating the cost-effectiveness of acupuncture therapy.

Methods The NICE OA economic analysis for acupuncture was replicated using ‘usual care’ (ie, no placebo/sham component) as the treatment comparator. A ‘transfer-to-utility’ technique was used to transform Western Ontario and McMaster Osteoarthritis scores into EQ-5D utility scores to allow quality-adjusted life year (QALY) gains to be estimated. QALY estimates were combined with direct incremental cost estimates of acupuncture treatment to determine incremental cost-effectiveness ratios (ICERs).

Results When ‘usual care’ was used as the treatment comparator, ICER point estimates were below £20 000 per QALY gained for each acupuncture trial analysed in the OA clinical guideline. In the original analysis, using placebo/sham acupuncture as the treatment comparator, ICERs were generally above £20 000 per QALY gained.

Conclusion The treatment comparator chosen in economic evaluations of acupuncture therapy is likely to be a strong determinant of the cost-effectiveness results. Different comparators used in the OA and LBP NICE guidelines may have led to the divergent recommendations in the guidelines.

INTRODUCTION

There has been considerable debate surrounding the recommendations made by the National Institute for Health and Clinical Excellence (NICE) regarding acupuncture therapy for low back pain (LBP) and osteoarthritis (OA).

The NICE OA guideline, issued in February 2008, recommended against the use of electroacupuncture and also concluded that existing evidence was not of sufficient consistency or strength to recommend the use of acupuncture.

Conversely, the NICE guidance for LBP, issued in May 2009, recommended that clinicians should consider offering a course of acupuncture therapy comprising a maximum of 10 sessions over a period of up to 12 weeks.

There were key differences in the economic evidence considered within the NICE OA and LBP guidelines. In the OA guideline, only non-UK economic evaluations were identified and these were deemed irrelevant for a National Health Service (NHS) context. Therefore, the guideline analysts conducted de novo economic analyses that were simplistic by necessity; time constraints dictated that full economic models could not be developed for all interventions considered by the guideline.

The economic analyses in the OA guideline (presented in appendix C of the guideline document) were conducted for acupuncture therapy, hyaluronan injections and various types of glucosamine and chondroitin. In order to attain a level of consistency throughout the guideline, all analyses compared the intervention in question to a placebo control group.

This was because trials of drug interventions in the context of OA typically compare the active agent against a ‘usual care’ control group that includes the administration of a placebo. The approach adopted in the guideline was considered to represent a suitable commonality for all analyses, rather than using a ‘no treatment’ or ‘usual care with no placebo’ comparator for the analysis of some interventions, and a placebo comparator for others.

In the LBP guideline, one relevant economic evaluation was identified, which examined the cost-effectiveness of acupuncture delivered by therapists trained in traditional Chinese medicine. The study found that acupuncture was associated with base case cost-effectiveness estimates (£4241 per quality adjusted life year (QALY) gained) that are acceptable to NICE;
acupuncture therapy was recommended by the guideline. Importantly, the treatment comparator in this single acupuncture trial was ‘usual care’, which did not involve a placebo component.

The issue of placebo control groups in the context of acupuncture therapy is a problematic concept. ‘Sham’ acupuncture is widely used in research settings but this could be argued to be more invasive than standard drug placebos. We do not seek to consider whether placebo or ‘usual care’ is the more appropriate comparator for an economic evaluation of acupuncture. In this paper, our specific objective is to address the inconsistency regarding treatment comparators across the economic analyses conducted in the NICE guidelines for OA and LBP, seeking to determine whether the choice of comparator is likely to have affected guideline recommendations.

METHODS

Methods adopted in the original NICE OA guideline

The NICE OA guideline adopted a simplistic form of evaluation to provide estimates of cost-effectiveness for a range of different interventions, including acupuncture. For each intervention, published papers that reported relevant randomised controlled trials (RCTs) were identified from a systematic review of the clinical evidence completed for the guideline; for each paper identified, the cost-effectiveness of the respective interventions was estimated. The adopted methodology for the economic analyses is described, briefly, below. Full details of the methods are described in appendix C of the NICE OA guideline.

An NHS payer perspective was taken for the analysis of costs, as required by NICE. However, only the direct costs of the interventions were considered, defined as the cost of physiotherapists’ time delivering treatment sessions and the needling costs for conducting acupuncture. It was assumed that each acupuncture session took 30 min. Unit costs were based on standard UK cost sources.

Due to the paucity of primary data sources for utility estimates in OA populations, transfer-to-utility (TTU) techniques have been derived to transform Western Ontario and McMaster Osteoarthritis (WOMAC) scores into utility estimates. Total WOMAC scores are made up of three subscales—pain (with scores between 0 and 20), function (with scores between 0 and 68) and stiffness (with scores between 0 and 8)—with an overall scoring range from 0 (no pain, stiffness or functional limitations) to 96 (severe pain, stiffness and functional limitations). WOMAC scores represent a disease-specific outcome measure for OA, and therefore are not suitable outcome measures for conventional economic analysis, where the objective is to provide evidence of cost-effectiveness that permits comparability across different disease areas. Such broad, generic comparisons are required in the context of NICE guidance because of the need to allocate healthcare resources across a multitude of competing demands.

Barton and colleagues have published a mapping algorithm that allows WOMAC scores to be transformed into EQ-5D utility scores based on a UK dataset, and it was this scoring procedure that was used in the NICE guideline. Mapping exercises to derive utility scores are open to error but they provide a pragmatic solution in scenarios where estimates from accepted utility instruments are not directly available. The EQ-5D is a generic instrument that measures health-related quality of life across five domains (mobility, self-care, usual activities, pain/discomfort, anxiety/depression), producing a single index value for health status. For UK-based technology appraisals, the EQ-5D is the preferred instrument for estimating utility scores. EQ-5D utility scores range from −0.594 to 1.000 and are interpreted on a 0 (a health state equivalent to death) to 1 (full health) scale; negative scores reflect particularly severe health states valued as worse than being dead.

Within each identified study, WOMAC scores measured over time were transformed into EQ-5D scores for the respective acupuncture and placebo groups. The QALY gain was then estimated for each group compared to their baseline utility. This allowed incremental cost-effectiveness ratios (ICERs) to be calculated using the following formula:

\[
\text{ICER} = \frac{(\text{cost of experimental intervention} - \text{cost of control intervention})}{(\text{QALY gain in the experimental intervention} - \text{QALY gain in the control intervention})}
\]

In the NICE OA guideline, studies were included in the economic analysis if they had an overall study sample size greater than 90, reported total WOMAC scores by treatment group, and included a placebo control group as the treatment comparator.

Methods adopted in this novel reanalysis

In our reanalysis of the economic evaluation of acupuncture therapy we used similar methods to those used in the original guideline, except that a ‘usual care’ comparator was used instead of a placebo control group.

The costs used in our analysis are presented in table 1. In the original NICE guideline, sham acupuncture treatments were assigned a zero cost because they were considered to be placebo controls rather than active interventions. In keeping with the simplistic approach adopted in the guideline, our reanalysis assumes that healthcare resource use outside of the acupuncture intervention itself is identical across treatment groups. Hence only the direct intervention costs associated with acupuncture therapy were included in the reanalysis (ie, by definition, this will result in a positive ‘incremental’ cost regarding acupuncture therapy). In line with the original NICE analysis, intervention costs were calculated under the assumption that each acupuncture session lasted 30 min. The cost per session was therefore £18 (0.5 multiplied by the hourly cost presented in table 1, plus £1 for the acupuncture needles). This cost was multiplied by the number of sessions in order to calculate total intervention costs.
We used Barton and colleagues’ ‘Model C’ procedure, which was identified as the appropriate scoring method in the absence of detailed information on the age and sex of participants. Equation (2) was used to transform total WOMAC scores into utility scores. This equation differs slightly from that used in the NICE OA guideline due to small alterations made by the authors after providing their data for use on the guideline and the subsequent publication of their research. Typesetter: In equation (2), change EQ - 5D to EQ-5D (i.e. remove spaces around hyphen)

Predicted EQ – 5D score =

$$\text{Predicted EQ – 5D score} = 0.74716 + (0.00081 \times \text{totalWOMAC}) + (-0.00012 \times \text{totalWOMAC}^2)$$

In our reanalysis, studies were only included if they met the criteria set out within the NICE OA guideline (summarised above) and if a ‘usual care’ arm that involved no acupuncture placebo was included in the trial. Using the methods described above, we replicated the analyses undertaken in the NICE OA guideline and estimated ICERs for acupuncture therapy compared to ‘usual care’. These results were then compared to the estimated ICERs when using a placebo-controlled comparator—that is, the method of analysis undertaken in the NICE OA guideline.

RESULTS

In the NICE OA guideline, three acupuncture studies and one electroacupuncture study were included. Of these studies, only the acupuncture studies are included in this reanalysis because the electroacupuncture trial did not include a ‘usual care’ arm without a placebo component.

The RCT reported by Berman et al compared 23 sessions of acupuncture over a 26-week period to sham acupuncture (placebo) and ‘usual care’ in patients with knee OA (n=570). WOMAC pain and function scores were measured at baseline, and 4, 8, 14 and 26 weeks. WOMAC stiffness scores were not reported. Accordingly, we estimated the total WOMAC score by transforming the WOMAC range of 0–88 (representing the range of scores relating to the pain and function subscales) into a score out of 96 by multiplying the 0–88 score by the factor 96/88 . This matches the approach taken in the NICE OA guideline.

Scharf et al reported an RCT in which acupuncture was compared to sham acupuncture (placebo) and ‘usual care’ in patients with knee OA (n=1007). The acupuncture regime involved 10 sessions over a 6-week period, with five additional sessions for patients who were perceived to be continuing to benefit from treatment at the end of the 6-week period. Total WOMAC scores were measured at baseline, and 13 and 26 weeks.

Witt et al reported results of an RCT that compared acupuncture to sham acupuncture and a waiting list group in patients with knee OA (n=294). Twelve sessions of acupuncture were administered over an 8-week period. The waiting list control group received no acupuncture for the initial 8 weeks of the trial, but then received the acupuncture intervention. Although total WOMAC scores were reported at baseline, and 8, 26 and 52 weeks, a direct comparison between acupuncture and ‘usual care’ can only be made over the first 8 weeks of the study, before the waiting list group received acupuncture.

The results of our reanalysis with ‘usual care’ (ie, no placebo component) as the treatment comparator are presented in table 2; results of the placebo comparison are presented in table 3. Note that the values in this table for the Berman et al and Scharf et al studies differ slightly from those in the NICE OA guideline, due to the small changes to the Barton et al algorithm. Results for the Witt et al study are substantially different because in our reanalysis we were only able to include an 8-week time-frame to ensure that directly comparable cost-effectiveness estimates for the ‘usual care’ and sham acupuncture comparisons could be obtained. In the NICE OA guideline, a 52-week time-frame was used and this led to increased QALY gain estimates because it appeared that incremental benefits continued to be accrued after treatment discontinuation at 8 weeks. Hence, it is possible that our cost-effectiveness estimates based on the Witt et al study are overestimates.

When the treatment comparator is ‘usual care’, the estimated ICER is below £20 000 per QALY gained in all three RCTs (see table 2). Conversely, when the treatment comparator was a placebo control group (ie, sham acupuncture, as in the original NICE OA guideline), the estimated ICER was above £20 000 in each of the RCTs.

DISCUSSION

Our analysis demonstrates that cost-effectiveness estimates for acupuncture therapy in the management of OA are likely to be heavily dependent on the treatment comparator used in the analysis. NICE typically considers interventions to be cost-effective if their ICER is below £20 000; our analysis demonstrates that if acupuncture had been compared to ‘usual care’ rather than placebo in the NICE OA guideline, it could have received a positive recommendation. It may be argued that the sole reason acupuncture received a positive recommendation in the LBP guideline, but not in the OA guideline, was that the economic analyses used ‘usual care’ as the treatment comparator.

Table 1 Unit costs of providing acupuncture treatment

<table>
<thead>
<tr>
<th>Resource use</th>
<th>Unit cost</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physiotherapist time in clinic</td>
<td>£34 per hour</td>
<td>Unit costs of health and social care, 2010</td>
</tr>
<tr>
<td>Acupuncture needles</td>
<td>£1 per session</td>
<td>Assumption based on the National Institute for Health and Clinical Excellence osteoarthritis guideline</td>
</tr>
</tbody>
</table>
As noted by the authors of the NICE OA guideline, it is important to maintain consistency in economic evaluation methodology in order to obtain directly comparable estimates of cost-effectiveness across the broad range of interventions evaluated in a single guideline. Within the OA guideline, most trials included a placebo version of the respective intervention as part of the control group. Therefore, when faced with a trial that includes both a placebo control group and a non-placebo control group, it may seem reasonable to use the placebo control group as the comparator in order to promote consistency.

In the LBP guideline it was stated that ‘seeing an acupuncturist was better than usual care but… there is not much difference between acupuncture and sham. However, sham acupuncture is used as an active form of treatment by some practitioners, therefore this should be considered as a possible treatment’ (p. 157). This highlights the nuances associated with placebo control groups and their potential effects. The authors of the LBP guideline appear to suggest that sham acupuncture should not be viewed in the same light as more conventional placebo interventions (eg, sugar pills in drug trials) because there is an evident, demonstrable active component. However, to suggest that a ‘sham’ intervention may be considered as a viable treatment option within the NHS would require further primary research. It is not appropriate to assume that trial results for sham acupuncture interventions would be replicated in routine clinical practice because recipients think they are receiving the full acupuncture intervention; reported health benefits may well be different if recipients know this is not the case.

One final important consideration is worthy of note. The NICE OA guideline stated that ‘there is not enough consistent evidence of clinical- or cost-effectiveness to allow a firm recommendation on the use of acupuncture for the treatment of osteoarthritis’. The guideline development group (GDG) may not have been convinced of the clinical effectiveness of acupuncture and, therefore, even if ‘usual care’ (however defined by the respective RCTs) was used as the treatment comparator for acupuncture in the economic analysis, the GDG may have decided not to recommend acupuncture for use in the NHS based on the clinical evidence available at the time.

### Study limitations

Inevitably, there are limitations to the simplistic cost-effectiveness analysis presented in this paper, and the comparable analysis reported in appendix C of the NICE OA guideline. Regarding the methods adopted in this paper and the OA guideline, these limitations are: (1) the economic analysis includes only the direct costs of providing acupuncture (no adverse events or other related healthcare resource use are considered); (2) health outcomes are not extrapolated to consider long-term effects (for example, only an 8-week follow-up is available for the study reported by Witt et al); and (3) the TTU analysis provides utility scores that are subject to significant uncertainties that are not characterised in the economic analysis (ie, no sensitivity analysis was undertaken).

These limitations were recognised by the NICE OA guideline authors, who were seeking to provide ‘ball-park’ estimates of cost-effectiveness for several interventions. The objective of this reanalysis was to explore the implications of using a different treatment comparator in the original NICE OA guideline analysis of acupuncture therapy. Accordingly, it was necessary to maintain methodological consistency with the guideline, and report point estimate...
Concluding comments
The treatment comparator included in economic evaluations of acupuncture therapy has important implications for the estimation of cost-effectiveness. Recommendations made in the NICE OA guideline regarding acupuncture therapy may have been different if ‘usual care’ with no placebo was used as the treatment comparator.

Contributors NRL: conception, design, analysis and interpretation of data, drafting of the manuscript, final approval given. ACB: analysis and interpretation of data, critical revision of manuscript, final approval given. NRL was a member of the National Institute for Health and Clinical Excellence Osteoarthritis Guideline Development Group.

Competing interests NRL was a member of the National Institute for Health and Clinical Excellence Osteoarthritis Guideline Development Group.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES
Inconsistencies in NICE guidance for acupuncture: reanalysis and discussion

Nicholas R Latimer, Aathish Chandrika Bhanu and David G T Whitehurst

Acupunct Med 2012 30: 182-186 originally published online July 3, 2012
doi: 10.1136/acupmed-2012-010152

Updated information and services can be found at:
http://aim.bmj.com/content/30/3/182

These include:

References
This article cites 13 articles, 5 of which you can access for free at:
http://aim.bmj.com/content/30/3/182#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections
Editor's choice (34)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/