Efficacy of dexmedetomidine compared with midazolam for sedation in adult intensive care patients: a systematic review

R. Adams, G. T. Brown, M. Davidson, E. Fisher, J. Mathisen, G. Thomson and N. R. Webster*

University of Aberdeen, Foresterhill, Aberdeen AB25 2ZD, UK
* Corresponding author. E-mail: n.r.webster@abdn.ac.uk

Summary. Patients in Intensive Care Unit (ICU) often require sedatives which commonly include midazolam and the more recently developed α2-receptor agonist, dexmedetomidine. It was our aim to compare the sedative and clinical effectiveness of dexmedetomidine vs midazolam in adults admitted to ICU, using an objective appraisal of randomized control trials. Medline, Embase, SCOPUS, Web of Knowledge, Cinthal, the United States National Library of Medicine, and the Cochrane Database of Systematic Reviews were searched using keywords: ‘dexmedetomidine’, ‘midazolam’, and ‘intensive care’. These were limited to human studies and adults (>18 yr old). Six randomized controlled trials were found and were critically appraised using a standardized appraisal method. Two papers described the time spent by each intervention group within a specified target sedation range and both found no statistically significant difference between midazolam and dexmedetomidine ($P = 0.18$ and $P = 0.15$). A third paper found no statistically significant difference in the length of time that patients were sedated within a target zone ($P = 0.445$). Two additional pilot studies did not report $P$ values as they were insufficiently statistically powered. A final paper found that, of the eight occasions measured, patients on dexmedetomidine were more often within the target sedation range than patients on midazolam. The sedative benefits of dexmedetomidine vs midazolam remain inconclusive. While some secondary outcomes showed clinical effectiveness of dexmedetomidine, more research is needed to validate the findings of these studies.

Keywords: dexmedetomidine; intensive care unit; midazolam; sedation

Patients admitted to the ICU are usually in need of invasive and uncomfortable interventions such as mechanical ventilation. To reduce anxiety, increase tolerance, and improve outcomes of such interventions, sedation is common practice.1 Traditionally, sedative agents administered in the ICU are γ-aminobutyric receptor agonists (GABA) which include the benzodiazepines (usually midazolam) and propofol.2 Optimum sedation is vital in striking a balance between providing pain relief and maintaining patient calm while preventing over-sedation and unnecessarily lengthy ICU stays.3 Many protocols advise daily sedation interruptions to assess the level of sedative in the patient and to avoid over-sedation.4

Dexmedetomidine has been studied as an alternative to traditional GABA-based sedation in the ICU. As a selective α2-receptor agonist, it acts at the locus coeruleus and spinal cord to exert anxiolytic and sedative effects without respiratory depression.5 Furthermore, there is evidence to suggest that administration of dexmedetomidine instead of standard sedatives (propofol or midazolam) in a critical care setting significantly reduces the incidence of delirium.6 The United States Federal Drug Administration has, however, advised that it is only used for short-term sedation (<24 h) because of adverse effects such as tachypnoea, complications of respiratory failure, acute respiratory distress syndrome, and agitation associated with longer administration times.7 Considering that the sedation needs of critically ill patients can often be for weeks at a time, this questions the suitability of dexmedetomidine as a sedative in the ICU setting.

Several sedation scoring scales have been developed for the assessment of sedation level and are used in studies to assess the amount of time a patient spends within a desirable ‘target range’. The Ramsey Sedation Scale (RSS) was the first standardized procedural measurement for sedation.8 The RSS scores patients between 1 and 6, with 1 corresponding to an anxious or agitated state and 6 to no response. The Riker Sedation and Agitation Score (RSAS) is similar and scores between...
1 for fully sedated and 7 for dangerously agitated. The Richmond Agitation–Sedation Scale (RASS) is a similar score which has been shown to correlate directly with other, more objective, measures of sedation such as the bispectral index (BIS). Haemodynamic variables such as heart rate (HR) and arterial pressure (AP) also provide objective measures by which sedation level can be assessed, although these are subject to other physiological factors.

We compared the sedative and clinical effectiveness of dexmedetomidine with midazolam, used in adults admitted to ICU using an objective appraisal of randomized control trials.

**Methods**

Medline (1946–present), Embase (Embase Classic + Embase, 1947–May 15, 2012), SCOPUS, Web of Knowledge, CINAHL and the United States National Library of Medicine were searched and also the Cochrane Database of Systematic Reviews (2005–April 2012). MeSH terms and keywords were defined as ‘dexmedetomidine’, ‘midazolam’, and ‘intensive care’, and these terms were combined (Fig. 1). For Medline and Embase, the search was then limited to studies that were randomized control trials involving humans and patients aged >18 yr old. Our search encompassed papers and conference abstracts in all languages. Full texts were then retrieved for those deemed suitable and these were again assessed for relevance. References in the retrieved articles were then scrutinized for any further studies. The papers were then critically appraised.

**Results**

Effectiveness of sedation was the main primary outcome in all six studies, though its definition varied slightly between them (Table 1). Jakob and colleagues, Riker and colleagues, Riker and colleagues, and Ruokonen and colleagues described effectiveness as the proportion of time spent within a target sedation range (or, as explained by Jakob; ‘maintaining sedation’). Senoglu and colleagues defined it as achieving a target level of sedation with respect to certain measurement scales (e.g. RSS, RSAS, and BIS), and in the study by Esmaoglu and colleagues, we believed it to be the duration of sedation; however, this was unclear. The pilot study by Riker and colleagues was presented at a conference and is only available as an abstract with limited results. The author was contacted and suggested that a full article publication was not possible as it was a pilot study with recruitment being ‘too small for any meaningful publication’.

The studies by Jakob and colleagues, Riker and colleagues, and Ruokonen and colleagues assessed patients using RASS although the target sedation ranges differed between them: Riker and colleagues used a range of −2 to 1 while Jakob and colleagues had a target range of −3 to 0. Ruokonen and colleagues determined target RASS range before starting the treatment, which varied from −4/−3 to 0. It cannot be ascertained from the paper how the target range was identified, but it appears to vary between study centres, although possibly between patients as well. Riker and colleagues found no statistically significant difference between treatment groups for time spent in target sedation range ($P=0.18$). Jakob and colleagues also found no statistically significant difference between their midazolam and dexmedetomidine treatment groups ($P=0.15$). Statistical significance for the primary outcome in the Ruokonen and colleagues study was not assessed with regards to midazolam as the comparator group was standard care [consisting of either propofol (28 patients) or midazolam (16 patients)]. The pilot study by Riker and colleagues states that the dexmedetomidine group spent 84.6% in the target sedation range compared with 77.2% for the midazolam group. However, no range is stated and no $P$-values were given.

Senoglu and colleagues found that patients’ average measurements for both the RSS and RSAS were within the target range on more occasions in the dexmedetomidine group than in the midazolam group. The target ranges of sedation in this study were from RSS 2 to 3 and between 3 and 4 using RSAS. Out of eight sedation level measurements $>24$ h (baseline, 1, 2, 4, 6, 8, 12, and 24 h), the dexmedetomidine group had an average of seven values within target range, compared with only five values from the midazolam group for RSS and 6 and 5, respectively, using RSAS. Both groups achieved BIS score of more than the target level of 85 throughout the duration of the study ($P<0.05$).

Esmaoğlu and colleagues target range of sedation was between RSS of 2 and 3 and there was no statistically significant difference between the groups in length of time sedated ($P=0.445$).

Length of ICU stay (LOS) was used as a secondary outcome in four of our six trials (Table 2). Only Esmaoglu and colleagues found a statistically significant difference in LOS between the two drugs ($P=0.021$). This was in favour of dexmedetomidine over midazolam with results of 45.5 and 83 h spent in ICU, respectively.

Haemodynamic characteristics were assessed as secondary outcomes for three of the six papers. Senoglu and colleagues measured the HR, AP, and arterial blood gases (ABG) of all patients during sedation. Unfortunately, the results for this study were illustrated in graphical form only and so could not be interpreted accurately. However, a lower rate for dexmedetomidine compared with midazolam was seen over the recorded 24 h period ($P=0.05$). In addition to the measurement of HR, Esmaoglu and colleagues also assessed mean arterial pressure (MAP). A statistically significant reduction in HR during the first 24 h was seen in the dexmedetomidine group ($P<0.05$), though no difference was found at 48 and 72 h. Esmaoglu and colleagues also found that patients who were sedated with dexmedetomidine had a lower MAP than those receiving midazolam between 3 and 24 h, though at 48 and 72 h the opposite was true. However, only the results from hours 5, 6, 12, and 24 h were statistically significant ($P<0.05$). Senoglu and colleagues reported that patients receiving dexmedetomidine had lower systolic and diastolic pressures during the first 2 h than patients who received midazolam ($P<0.05$). There was no statistically significant
difference in AP at 4 h. ABGs were measured only in the Senoglu study, which found that in both groups $P_{aCO_2}$ decreased ($P<0.05$), there was no difference in $P_{aO_2}$. The pilot study by Riker and colleagues noted that both study groups had decreases in HR, with the dexmedetomidine group showing a greater decrease (12.7 vs 2.0 bpm). Systolic AP was preserved in both groups, ↑7.1 (13.0) in the dexmedetomidine vs ↓0.6 (14.0) in the midazolam group) despite baseline systolic AP being lower in the dexmedetomidine group [105 (13) vs 125 (21) midazolam, $P<0.01$].

Some of the papers included additional outcomes. Jakob and colleagues$^{15}$ and Ruokonen and colleagues$^{18}$ compared nurses’ assessment of patient ability to cooperate with care, level of arousal, and ability to communicate pain using a visual analogue scale. Both papers concurred that patients receiving dexmedetomidine were more rousable, cooperative and able to communicate pain ($P<0.001$). Riker and colleagues$^{16}$ and Ruokonen and colleagues$^{18}$ compared prevalence of and duration without delirium. Riker and colleagues found that patients receiving dexmedetomidine had lower
Table 1  Characteristics of the selected studies. RSS, Ramsey Sedation Score; RASS, Richmond Agitation-Sedation Score; RSAS, Riker Sedation Agitation Score; BIS, bispectral index; ICU, Intensive Care Unit; N/A, not applicable

<table>
<thead>
<tr>
<th>Study</th>
<th>Dexmedetomidine</th>
<th>Midazolam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Esmaoglu and colleagues, 2009</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Primary outcome(s)</td>
<td>Effectiveness of sedation—no definition provided</td>
<td></td>
</tr>
<tr>
<td>Primary outcome measurements</td>
<td>RSS</td>
<td>RSS</td>
</tr>
<tr>
<td>Target sedation range</td>
<td>RSS: 2 to 3</td>
<td>RSS: −3 to 0</td>
</tr>
<tr>
<td>% Time at target sedation and P-value</td>
<td>Unclear</td>
<td>60.7</td>
</tr>
<tr>
<td>Main findings</td>
<td>Dexmedetomidine reduces HR, need for anti-hypertensives and ICU length of stay compared with midazolam</td>
<td>Dexmedetomidine is non-inferior to midazolam for ICU sedation in adults</td>
</tr>
<tr>
<td>Comments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Riker and colleagues, 16 2009</td>
<td>244</td>
<td>122</td>
</tr>
<tr>
<td>Primary outcome(s)</td>
<td>Proportion of time spent at target sedation level</td>
<td>Proportion of time spent at target sedation level</td>
</tr>
<tr>
<td>Primary outcome measurements</td>
<td>RASS</td>
<td>RASS</td>
</tr>
<tr>
<td>Target sedation range</td>
<td>RASS: −2 to 1</td>
<td>RASS: −4 to 3 to 0</td>
</tr>
<tr>
<td>% Time at target sedation and P-value</td>
<td>77.3</td>
<td>64</td>
</tr>
<tr>
<td>Main findings</td>
<td>No difference in time spent within target RASS range. Lower incidence of delirium (and duration)</td>
<td>Non-inferiority of dexmedetomidine to standard care was not confirmed</td>
</tr>
<tr>
<td>Comments</td>
<td>This study was funded by Hospira, Inc., Lake Forest, IL, USA, which manufactures dexmedetomidine</td>
<td>This study was sponsored by Orion Pharma, Helsinki, Finland, through research contracts negotiated directly with the participating hospitals</td>
</tr>
</tbody>
</table>

Table 2  The median length of time spent in ICU (h).*Mann–Whitney t-test, **Gehan-Wilcoxon test, and ***log-rank test

<table>
<thead>
<tr>
<th>Study</th>
<th>Dexmedetomidine</th>
<th>Midazolam</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Esmaoglu and colleagues20 (range)</td>
<td>45.5 (15–118)</td>
<td>83 (15–312)</td>
<td>0.02*</td>
</tr>
<tr>
<td>Jakob and colleagues15 (IQR)</td>
<td>211 (115–831)</td>
<td>243 (140–630)</td>
<td>0.27**</td>
</tr>
<tr>
<td>Riker and colleagues16 (95% CI)</td>
<td>141.6 (136.8–168)</td>
<td>182.4 (160.8–206.4)</td>
<td>0.24***</td>
</tr>
<tr>
<td>Ruokonen and colleagues18 (range)</td>
<td>5.5 (1.7–19.5)</td>
<td>5.8 (standard care—5.7) (1.8–29.0)</td>
<td>N/A (0.411)</td>
</tr>
</tbody>
</table>
incidences of delirium and on average had more delirium-free days \( (P<0.001 \text{ and } P=0.002, \text{ respectively}) \). There was an increased use of open label midazolam in those treated with dexmedetomidine \( (P=0.02) \) but no difference between groups for fentanyl use \( (P=0.25) \). Ruokonen and colleagues \( ^{18} \) found that delirium was more common in the dexmedetomidine group \( (43.9\%) \) vs the standard care (propofol and midazolam) group \( (25\%) \) and this result was significant \( (P=0.035) \). Jakob and colleagues found that the median duration of mechanical ventilation for patients on dexmedetomidine was less \( (123 \text{ h compared with } 164 \text{ h for those on midazolam}) \) \( (P=0.03) \). Ruokonen and colleagues \( ^{18} \) also investigated the median duration of mechanical ventilation, though again this was compared with ‘standard care’ and not midazolam. The median duration of mechanical ventilation was 77.2 h \( (17.5–338.8 \text{ h}) \) in the dexmedetomidine group and 110.6 h \( (20.1–675.0 \text{ h}) \) in the standard care group \( (P=0.109) \). Riker and colleagues compared time until extubation and also found that this was less for dexmedetomidine \( (\text{median time } 3.7 \text{ vs } 5.6 \text{ days}) \) \( (P=0.01) \).

The different sources of potential bias have been outlined in Table 3 and an assessment for the overall risk of bias for each paper has been made.

### Discussion

We could find no ideal study comparing the effectiveness of dexmedetomidine with other sedative agents commonly used on ICU. The main problem with assessing the effectiveness of sedation is that most measurements are made from subjective scales. Only one of the studies included BIS as an objective measurement. \( ^{19} \) Precise measurements of target sedation are therefore made subjectively with an inherent risk of interpretation bias and inconsistency. Only in one study were RSS and RSAS assessments made by the same blinded investigator to minimize inconsistency. \( ^{19} \)

In Esmaoglu and colleagues, \( ^{20} \) RSS was used to ensure that patients were at an appropriate sedation level, but any record of how long patients were maintained at these target levels \( (2–3 \text{ on the RSS}) \) was omitted. ‘Effectiveness’ of sedation perhaps meant the duration of sedation in general, which provided unreliable results as additional propofol was administered if ‘sedation became inadequate (RSS < 2)’, and more patients in the midazolam group received additional propofol than in the dexmedetomidine group \( (12 \text{ vs } 9) \). Thus, the primary outcome for this study did not present usable comparable results and was ignored in both the discussion and conclusion. In the study by Jakob and colleagues, \( ^{15} \) patients who were given propofol were not analysed independently from those who did not receive propofol—a gain leading to difficulties in accurately interpreting the conclusions.

It would seem that dexmedetomidine may reduce the length of ICU stay, though only one \( ^{20} \) recorded a statistically significant difference between treatment groups. This result concurs with previous reviews similar to this topic. The meta-analysis by Tan and colleagues \( ^{21} \) notes that dexmedetomidine may reduce the length of ICU stay, though because of the limited evidence available, this is uncertain.

Furthermore, there are confounding issues to consider. Two studies \( ^{16} \) \( ^{19} \) seemed to include similar patients, while Esmaoglu and colleagues \( ^{20} \) had a very specific population (solely eclamptic patients) and this is possibly why it was the only trial to report a statistically significant result. Another confounder for this outcome was found in the trial by Jakob and colleagues \( ^{15} \) where the study drug was only administered between ICU days 3 and 14, though mechanical ventilation was considered in patients for up to 45 days. Before Day 3 and after Day 14, standard care was administered which was not defined. Within each of the study groups, the patients were comparable at baseline after randomization, though between studies they were not, because of different inclusion and exclusion criteria. The results between studies may not be directly comparable and this could account for differences in LOS.

---

**Table 3** Assessment of Methodological Quality of Trials: risk of bias. Although it is good research practice to adopt a double-blind trial design, it would be unfair and inaccurate to say that a lack of patient blinding in the particular setting of these trials would introduce a high risk of bias because the primary outcome assesses the level of sedation which is something that the patients themselves have no manual control over.

<table>
<thead>
<tr>
<th>Author</th>
<th>Randomization process</th>
<th>Allocation concealment</th>
<th>Incomplete outcome data</th>
<th>Selective reporting bias</th>
<th>Free of other bias</th>
<th>Overall risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Esmaoglu and colleagues (^{20})</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Jakob and colleagues (^{15})</td>
<td>Low</td>
<td>Unclear</td>
<td>High</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Unclear</td>
</tr>
<tr>
<td>Riker and colleagues (^{17})</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Riker and colleagues (^{16})</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Ruokonen and colleagues (^{18})</td>
<td>Unclear</td>
<td>Unclear</td>
<td>High</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Unclear</td>
</tr>
<tr>
<td>Senoglu and colleagues (^{19})</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>
A sedative agent that can provide cardiovascular stability may be expected to be beneficial to patients in ICU. Patients who were sedated with dexmedetomidine in a trial by Frolich and colleagues support the decrease in AP reported in the trials by Esmaoglu and colleagues and Senoglu and colleagues. A study by Ickeringill and colleagues also found that dexmedetomidine decreased mean HR and mean systolic AP.

Jakob and colleagues concluded that patients given dexmedetomidine were more rousable, cooperative and able to communicate pain. Ruokonen and colleagues concurred with this finding, although this was compared with standard care and not midazolam alone. Riker and colleagues found a lower incidence of delirium and longer delirium-free duration in those on dexmedetomidine. However, they only looked for delirium while patients were taking the study drug and 48 h after cessation. Ruokonen and colleagues, again comparing with standard care, found that delirium was more prevalent in the dexmedetomidine group. In the same paper, it was concluded that there was an equal need for the use of open label midazolam and fentanyl in patients in either treatment group. Though perhaps true, this may have confounded the primary outcome as the patients in the dexmedetomidine group were given midazolam and were not separated from those who did not receive it in the results. As a secondary outcome, Jakob and colleagues found that patients in the dexmedetomidine group received mechanical ventilation for less time than patients in the midazolam group. This was also found in the Ruokonen and colleagues group, although was again compared with a standard care group. Riker and colleagues also found that patients were extubated sooner in the dexmedetomidine group. This paper states that patients who were not extubated, were not included in the analyses. It also fails to state whether patients who were no longer on their study drug and were extubated, were included in the analysis. However, patients in the latter study were allowed the longer maximum duration of study drug administration of 30 days.

**Limitations of the studies**

The six studies are clear in stating aims comparing the sedative efficacy of dexmedetomidine against that of midazolam in an ICU setting using adult patients as participants.

However, the potential for confounding because of underlying disease is not clearly outlined. The international recruitment of patients seen in Riker and colleagues could also lead to problems with treatment standardization. This was compensated for with a start-up meeting for all investigators and research coordinators. Here, they were trained in using the RASS and how to titrate the blinded study drug. The use of opiates vs sedation was also standardized.

Having failed to provide an adequate definition of the primary outcome of ‘sedation effectiveness’, Esmaoglu and colleagues struggled to achieve their study aim and reach reliable conclusions. Jakob and colleagues, while using a recognized and validated system (RASS), do not mention blinding or single observer measurement. Senoglu and colleagues reduced bias by using the same blinded investigator and BIS, although others have questioned the objectivity of this measurement. Esmaoglu and colleagues used RSS, but the lack of blinding by assessors must introduce bias.

As non-inferiority studies, Jakob and colleagues and Ruokonen and colleagues are not fully assessing dexmedetomidine’s clinical effectiveness. Senoglu and colleagues provide a full description of the drug regimen, although within the discussion there is confusion between results over which agent is more clinically effective.

Randomization methods varied between the studies. Esmaoglu and colleagues used a coin toss, Senoglu and colleagues used a computer-generated randomization schedule, while Jakob and colleagues and Riker and colleagues both used a central interactive voice system (Jakob and colleagues on a 1:1 basis stratified for study centre in blocks of 4, Riker and colleagues on a 2:1 basis, dexmedetomidine:midazolam). Randomization was claimed in the Ruokonen and colleagues study and stratification details were given, though there was no explanation as to how it was actually carried out.

Numerical inconsistencies were found in Jakob and colleagues where 7800 patients were stated as excluded though this was verified as 8311. Furthermore, 60 patients in the dexmedetomidine group and 51 patients in the midazolam group had treatment withdrawn, though this was actually found to be 64 and 52, respectively. Inconsistencies were also identified in the Ruokonen and colleagues paper, with 47 patients experiencing serious adverse events in the standard care group of 44 patients.

Esmaoglu and colleagues did not report side-effects. Senoglu and colleagues reported only one side-effect with one patient randomized to the midazolam group having treatment stopped because of oversedation. In both Riker and colleagues and Jakob and colleagues, side-effects were adequately reported and analysed.

The trial by Ruokonen and colleagues had a very limited number of patients and was not powered to an appropriate statistical level. This severely limits the ability to draw useful conclusions from the stated results.

**Application of the results to ICU practice**

The lower sample variance seen in large multi-centre trials such as Riker and colleagues and Jakob and colleagues mean that results are more likely to be representative of the true population. Smaller trials with more strictly selected patient populations such as Esmaoglu and colleagues and Senoglu and colleagues are less likely to be generally applicable to clinical practice. It is worth noting that the extensive exclusion criteria seen in all papers reduced the applicability of the findings of this review. A power calculation for the primary outcome was not performed in the pilot study by Ruokonen and colleagues, therefore, very little can be drawn from its results.

The ideal method for comparing dexmedetomidine with midazolam would be a double-blinded, randomized control
trial with a study sample of sufficient size to ensure statistically significant results. We believe that this has still to be performed and we have been unable to undertake a power calculation for such a study because of a lack of variance data. The ideal assessment method would be objective and easily replicated with high sensitivity and specificity. We feel that such an assessment method does not yet exist.

In conclusion, we have carried out a systematic review comparing the sedative qualities of dexmedetomidine and midazolam for adult patients in ICU. Six randomized control trials were selected and critically appraised. Overall, evidence for the sedative superiority of dexmedetomidine over midazolam remains inconclusive and highlights the need for further, more rigorously designed trials. However, based on the literature available, dexmedetomidine appears to be a safe alternative to midazolam and may be more cost-effective.

Authors’ contributions
R.A., G.T.B., M.D., E.F., J.M., and G.T. should all be considered first authors: they undertook the search and wrote the initial manuscript. N.R.W. provided the initial idea, supervision and helped with manuscript preparation.

Declaration of interest
None declared.

Funding
This work was supported by departmental funding only.

References

Appendix 1. Critical Appraisal of Clinical Trials

Initial points
How many patients?
What is the primary outcome?
Are there any secondary outcomes?
Were treatments randomly allocated?
Were all the patients accounted for? Lost from the follow-up?
Were outcomes assessed blind?
Design

Are the aims clearly stated? Why was the study carried out?
Was the sample size justified?

Formal sample size calculation carried out and detailed in the methods.
Discussion—what size of effect did the study have the power to detect?

Are the measurements likely to be valid and reliable?

Methods of measurement should be described in detail. Some effort must be made to standardize methods in multi-centre trials. Discuss how validity and reliability are assessed.

Could the choice of subjects influence the size of treatment effect?

The setting that patients were recruited from.
Diagnostic criteria for entry to trial.
Factors for exclusion from study.
Description of duration and severity of disease at entry to study.

Were there ambiguities in the description of the treatment and its administration?

Are the statistical methods described?

All statistical methods should be described and referenced in the methods sections. All statistical methods make some assumptions and it is encouraging if this is addressed. Lots of tests and exotic complicated statistical tests could suggest these were chosen because of the $P$-value they yielded. Simple methods should always be shown and compared with more complex ones.
Could the lack of blinding introduce bias?
Who was blinded and how was it done? How were treatments allocated?
Are the outcomes clinically relevant?

Conduct

Did untoward events occur in the study?

Initial designs can sometimes be hard to follow, subjects may not be contactable and others may disappear. These issues should be identified and dealt with in pilot studies. Occurrence in the main study may suggest inadequate preparation. Some untoward events can be entirely unpredictable but others may suggest the study is of poor quality.

How was the randomization carried out?

Analysis

Were the treatment groups comparable at baseline?
Were results analysed by intention to treat?
Was the statistical significance assessed?

The results of all research studies are influenced by the play of chance. Sometimes chance effects can appear quite large, especially when the sample size is small. Thus, the statistical significance of the main findings should be assessed. A $P$-value of $<0.05$ provides good evidence that the result is likely to be real rather than chance. Even smaller $P$-values such as $<0.01$ or below give extra confidence that the result was not a chance event. Confidence intervals can also be used and these provide extra information as to where the correct value is likely to lie. Tight confidence intervals are good, whereas a large range calls the size of effect into question.

Were the basic data adequately described?

Basic data should be given a mean/median and a standard deviation/inter-quartile range to help generalization to the reader’s patient population.

Do the numbers add up?

Inconsistencies should be explained. Failure indicates sloppiness. Large discrepancies are particularly hazardous.

Were side-effects reported?

Interpretation

What do the main findings mean?

Should be your own interpretation and not just accepting the conclusion given by the authors at face value.

How are null findings interpreted?

These can arise because of study design or size and should be interpreted carefully.

Are important effects overlooked?

Everything should be considered not just the facts that the author wants to portray.

How do the results compare with previous reports?

What implications does the study have for your practice?

How big is the effect and is it clinically important? Then assess overall quality of the study and then see if the circumstances and subject background is similar to local patients.

Handling editor: R. P. Mahajan