A randomized, controlled comparison of the efficacy and tolerability of low and high doses of haloperidol in the treatment of first-episode psychosis

Piet Oosthuizen, Robin Emsley, H. Jadri Turner and Natasha Keyter

Department of Psychiatry, University of Stellenbosch, Cape Town, South Africa

Abstract

While haloperidol is still widely used in the treatment of psychoses, the optimal daily dose remains a topic of controversy, particularly in first-episode psychosis. Previous studies have suggested that doses as low as 2 mg/d may be effective, whereas others have indicated superiority for higher over lower doses. This double-blinded, randomized controlled study compared the efficacy and tolerability of 2 vs. 8 mg/d of haloperidol over 6 wk in 40 subjects with first-episode psychosis. Both treatments were equally effective in reducing the PANSS Total and subscale scores. The low dose of haloperidol was better tolerated, with fewer extrapyramidal side-effects, less frequent use of anticholinergic medication and smaller elevations in prolactin levels. Using a low dose of haloperidol is at least as effective as, and better tolerated than a high dose of haloperidol in the treatment of first-episode psychosis.

Received 3 June 2003; Reviewed 1 July 2003; Revised 17 August 2003; Accepted 25 August 2003

Key words: First-episode psychosis, haloperidol, schizophrenia, treatment.

Introduction

While second-generation antipsychotics are increasingly being used for the treatment of psychosis, particularly in North America and Europe, haloperidol remains a very important drug throughout the rest of the world. This is mainly due to its widespread, off-patent availability and low acquisition cost (Emsley et al., 1999). Furthermore, there is considerable disagreement regarding the safety and efficacy of haloperidol compared to second-generation antipsychotics. On the one hand, some suggest that second-generation antipsychotics may be no more effective and only marginally better tolerated than haloperidol (Geddes et al., 2000), when the effect of an overly high dose of haloperidol is controlled for, whereas others believe that clear-cut efficacy and tolerability differences exist between conventional and second-generation antipsychotics (Davis et al., 2003). A recent Cochrane Review concluded that, although quite effective, haloperidol is so toxic that ‘given the choice, people with schizophrenia may wish to start another antipsychotic with less likelihood of causing parkinsonism, akathisia and acute dystonias’ (Joy et al., 2002).

In spite of its widespread use in clinical and research settings, there is a dearth of information regarding the optimal dose of haloperidol. Although the American Psychiatric Association (APA, 1997) and the PORT Treatment Recommendations (Lehman et al., 1998) advocate haloperidol doses of between 6 and 20 mg/d, there is now a considerable literature suggesting that the use of lower doses is associated with a more favourable side-effect profile, without loss of efficacy (Emsley and the Risperidone Working Group, 1999; Kapur et al., 1996; McEvoy et al., 1991; Zhang-Wong et al., 1999). Functional brain-imaging studies with haloperidol over the past decade have shown a clear-cut relationship between striatal D2 receptor occupancy on the one hand, and antipsychotic efficacy, hyperprolactinemia and the emergence of extrapyramidal side-effects (EPS) on the other (Farde et al., 1992). It has been suggested that there is a therapeutic window for D2 receptor occupancy, with antipsychotic response starting around the 60% occupancy level and EPS
developing at receptor occupancy levels around 80% (Kapur et al., 2000). This implies that careful titration of the haloperidol dose may achieve an antipsychotic effect without the emergence of untoward side-effects. Although this has never been explored in large-scale studies, the seminal paper on neuroleptic threshold with haloperidol by McEvoy et al. (1991) suggests that this may very well be the case.

Determining the optimal antipsychotic dose is of particular importance in first-episode psychosis, as these patients are much more sensitive to the effects of these compounds (McEvoy et al., 1991; Wolkin et al., 1989). We previously reported the results of a study indicating that very low doses of haloperidol are effective and well tolerated in subjects with first-episode psychosis (Oosthuizen et al., 2001). However, as with the study of McEvoy et al. (1991), it was open label and did not have a control arm. The question as to whether very low doses of haloperidol are as effective and better tolerated than standard doses has yet to be answered. The purpose of the present study was to further explore this issue by using a double-blind, controlled design.

Method

Forty subjects with a first episode of psychosis were included in the study. Subjects were recruited from the Tygerberg–Stikland academic complex of hospitals. Both of these hospitals are affiliated to the University of Stellenbosch in Cape Town, South Africa. The study was approved by the Ethics Committee of the University of Stellenbosch. Subjects provided written, informed consent before any study-related procedures were undertaken.

Inclusion criteria were: in-patients or outpatients, aged between 16 and 55 yr (extremes included) with a DSM-IV diagnosis of schizoaffective disorder, schizophrenia or schizoaffective disorder and a lifetime neuroleptic exposure of 4 wk or less. Exclusion criteria were a DSM-IV Axis I diagnosis other than schizophrenia or schizoaffective disorder, schizophrenia or schizo-affective disorder, alcohol or drug dependence (although alcohol and cannabis use were not exclusion criteria), depot antipsychotic treatment, significant medical illness and mental retardation. All subjects were evaluated by P.O. or H.J.T., who participated in regular inter-rater reliability training sessions and a high level of inter-rater reliability was maintained.

At baseline data were obtained on demographic information, personal, psychiatric, medical and family history, previous treatment and the use of concomitant medication. A complete physical, including neurological, examination was performed. The following rating scales were administered at baseline: Structured Clinical Interview for DSM-IV (SCID; First et al., 1994), Positive and Negative Symptom Scale (PANSS; Kay et al., 1987), Clinical Global Impression (CGI; Guy, 1976), Calgary Depression Rating Scale (Addington et al., 1993) and Extrapyramidal Symptom Rating Scale (ESRS; Chouinard et al., 1980).

As the mean end-point daily dose in the acute phase of our open-label study was approx. 1.78 mg/d (Oosthuizen et al., 2001) we chose 2 mg/d as the low dose arm for the study group. Haloperidol (8 mg/d) was used as the high dose. This was decided on the basis of previous studies suggesting that doses higher than 10 mg/d may be detrimental (Van Putten et al., 1992) and the only ‘dose-finding’ study with haloperidol, suggesting 8 mg/d to be the optimal dose (Zimbroff et al., 1997).

Subjects were randomized 1:1 into one of the two treatment groups, and received a single, identical-looking capsule per day. Capsules for the low-dose group contained 2 mg haloperidol, and for the high-dose group 2 mg/d for days 1–2, 4 mg/d for days 3–4, 6 mg/d for days 5–6, and 8 mg/d for the rest of the trial. Lorazepam was permitted for restless or agitated behaviour throughout the study. Orphenadrine was permitted for treatment-emergent EPS. All concomitant medications that were prescribed were recorded. The study duration was 6 wk. Subjects were assessed at weekly intervals with the PANSS, CGI, CGI-CIS, Calgary Depression Rating Scale and ESRS. Blood samples for prolactin levels were taken at baseline and at end-point.

Statistical analysis

Statistical analysis was performed with the help of the software package STATISTICA version 6 (Statsoft Inc., Tulsa, OK, USA). Analysis was conducted on the intent-to-treat population, with last observation carried forward (LOCF). Categorical variables were compared using \( \chi^2 \) or Fisher’s exact test (two-tailed in all cases), depending on expected frequencies. All numerical variables were first tested for normality of distribution using the Kolmogorov–Smirnov method. In cases where unrelated groups were compared in terms of numerical variables, we used either Student’s \( t \) test (parametric) or the Mann–Whitney \( U \) test (non-parametric). For correlations between pairs of numeric variables, we used either the Pearson product moment correlation coefficient (parametric), or the Spearman rank-order correlation coefficient (non-parametric). A significance level of 0.05 was used throughout.
Results

Forty subjects were included in the study, 20 in each treatment arm. The sample size was based on the previous study by Zhang-Wong et al. (1999).

The two groups were similar with regard to various demographic and clinical characteristics (Table 1). Of the 40 subjects randomized to the trial, 11 (27.5%) did not complete the full 6 wk of the study. Three of the non-completers were in the low-dose group (2 mg/d), whereas eight were in the high-dose group (8 mg/d). This difference did not, however, reach statistical significance ($p = 0.08$, Fisher’s exact test, two-tailed).

Reasons for subjects not completing the study were, for the high-dose and low-dose groups respectively, treatment-emergent EPS ($n = 6$ vs. $n = 0$) ($p = 0.02$, Fisher’s exact test, two-tailed); poor response ($n = 2$ vs. $n = 1$); patient withdrawal of consent ($n = 0$ vs. $n = 1$); and lost to follow-up ($n = 1$ vs. $n = 0$).

Efficacy

Primary measure

Both treatment arms showed a robust reduction in PANSS Total scores, and there were no significant differences between the groups at any stage of the trial (Tables 2 and 3). Also, there were no differences between the groups regarding PANSS Positive and General Psychopathology subscale scores (Tables 2 and 3). For the PANSS Negative subscale scores, there was a trend towards greater improvement in the low-dose group, although the difference only reached statistical significance at week 2 (Table 4). Mean CGI ratings at end-point were similar for the two groups (high-dose group $3.35 \pm 1.70$ vs. low-dose group $3.47 \pm 1.71$; $t = 0.23$, d.f. = 37, $p = 0.82$) as were depressive symptoms, as measured by the Calgary Depression Rating Scale (high-dose group $1.2 \pm 2.97$ vs. low-dose group $0.6 \pm 1.35$; $t = -0.8$, d.f. = 38, $p = 0.42$).

In view of the fact that some measures, particularly the change in negative symptoms, showed trends towards group differences without reaching statistical significance, we performed a post-hoc power analysis of the sample. We found that the power of the current sample size to predict a statistically significant difference between the two means for improvement in negative symptoms was only 0.25.

Due to the high dropout rate in the 8-mg group, we also conducted a completers-only analysis of the primary and secondary efficacy outcome measures. The results were essentially the same as the LOCF analysis.

Side-effects

Global EPS as assessed by the ESRS total score (items 1–55) showed highly significant differences between the two groups (Figure 1). While parkinsonism (ESRS Parkinsonism subscale) scores showed a significant change from baseline to end-point for both the 2-mg

Table 1. Demographic and clinical details of the two treatment groups (2 vs. 8 mg/d) at baseline

<table>
<thead>
<tr>
<th>Parameter (baseline)</th>
<th>2-mg group</th>
<th>8-mg group</th>
<th>$\chi^2$</th>
<th>$t$, d.f.</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (male:female)</td>
<td>14:6</td>
<td>10:10</td>
<td>1.7</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>Ethnic group</td>
<td>17:2:1</td>
<td>14:2:4</td>
<td>2.1</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>Marital status (single/married)</td>
<td>1:17</td>
<td>5:13</td>
<td></td>
<td>0.18*</td>
<td></td>
</tr>
<tr>
<td>Age (yr)</td>
<td>26.8 (± 7.3)</td>
<td>28.9 (± 8.1)</td>
<td>0.8, 34</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>PANSS Positive Scale</td>
<td>27.0 (± 5.7)</td>
<td>27.3 (± 3.9)</td>
<td>–0.19, 38</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td>PANSS Negative Scale</td>
<td>25.6 (± 10.0)</td>
<td>23.9 (± 6.4)</td>
<td>0.64, 38</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>PANSS General Scale</td>
<td>47.6 (± 9.5)</td>
<td>44.6 (± 8.2)</td>
<td>1.05, 38</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>PANSS Total score</td>
<td>100.1 (± 19.3)</td>
<td>95.7 (± 13.0)</td>
<td>0.84, 3.0</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>ESRS score</td>
<td>2.4 (± 4.8)</td>
<td>0.7 (± 1.5)</td>
<td>1.56, 38</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>CGI</td>
<td>5.4 (± 0.9)</td>
<td>5.2 (± 0.7)</td>
<td>0.80, 38</td>
<td>0.43</td>
<td></td>
</tr>
</tbody>
</table>

* Fisher’s exact test, two tailed.

Table 2. Percentage changes in PANSS Total and subscale scores from baseline to end-point (LOCF)

<table>
<thead>
<tr>
<th>PANSS subscale</th>
<th>Total group</th>
<th>2-mg group</th>
<th>8-mg group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>37.3 (± 26.9)</td>
<td>36.0 (± 27.6)</td>
<td>38.6 (± 26.8)</td>
</tr>
<tr>
<td>Negative</td>
<td>15.2 (± 31.4)</td>
<td>21.8 (± 31.0)</td>
<td>8.9 (± 31.3)</td>
</tr>
<tr>
<td>General</td>
<td>20.1 (± 38.4)</td>
<td>27.2 (± 22.9)</td>
<td>13.4 (± 48.6)</td>
</tr>
<tr>
<td>Total</td>
<td>24.7 (± 29.1)</td>
<td>28.9 (± 23.9)</td>
<td>20.7 (± 33.4)</td>
</tr>
</tbody>
</table>
The 2-mg group had significantly lower prolactin levels compared to the 8-mg group at the end-point (t = 3.49, d.f. = 10, p = 0.006) (Table 3). The total dose of orphendrine used in the 8-mg group was significantly higher than that in the 2-mg group (t = 2.5, d.f. = 37, p = 0.016).

Concomitant medications

The amount of lorazepam used did not differ significantly between the two groups, (48.7 ± 55.2 mg for the 2-mg group vs. 49.0 ± 55.2 mg for the 8-mg group) (t = 0.01, d.f. = 37, p = 0.99). However, the mean total dose of orphendrine used in the 8-mg group (1967.5 ± 2261.9 mg) was significantly higher than that in the 2-mg group (555.3 ± 943.5 mg) (t = 2.5, d.f. = 37, p = 0.016).

Four subjects (10%) received concomitant treatment with fluoxetine during the 6 wk. Three were from the low-dose group and one from the 8-mg group (p = 0.9, Fisher’s exact test, two-tailed).

Prolactin levels

There was a significant increase in prolactin levels from 14.1 ± 9.5 ng/ml at baseline to 51.4 ± 34.8 ng/ml at end-point (t = 3.49, d.f. = 10, p = 0.006) for the 8-mg group, but not for the 2-mg group, where mean prolactin levels only increased marginally from 14.1 ± 9.5 ng/ml to 14.8 ± 9.7 ng/ml (p = 0.15, paired t test).

Concomitant medications

The amount of lorazepam used did not differ significantly between the two groups, (48.7 ± 55.2 mg for the 2-mg group vs. 49.0 ± 55.2 mg for the 8-mg group) (t = 0.01, d.f. = 37, p = 0.99). However, the mean total dose of orphendrine used in the 8-mg group (1967.5 ± 2261.9 mg) was significantly higher than that in the 2-mg group (555.3 ± 943.5 mg) (t = 2.5, d.f. = 37, p = 0.016).

Four subjects (10%) received concomitant treatment with fluoxetine during the 6 wk. Three were from the low-dose group and one from the 8-mg group (p = 0.9, Fisher’s exact test, two-tailed).

Prolactin levels

There was a significant increase in prolactin levels from 14.1 ± 9.5 ng/ml at baseline to 51.4 ± 34.8 ng/ml at end-point (t = 3.49, d.f. = 10, p = 0.006) for the 8-mg group, but not for the 2-mg group, where mean prolactin levels only increased marginally from 14.1 ± 9.5 ng/ml to 14.8 ± 9.7 ng/ml (p = 0.15, paired t test).

Concomitant medications

The amount of lorazepam used did not differ significantly between the two groups, (48.7 ± 55.2 mg for the 2-mg group vs. 49.0 ± 55.2 mg for the 8-mg group) (t = 0.01, d.f. = 37, p = 0.99). However, the mean total dose of orphendrine used in the 8-mg group (1967.5 ± 2261.9 mg) was significantly higher than that in the 2-mg group (555.3 ± 943.5 mg) (t = 2.5, d.f. = 37, p = 0.016).

Four subjects (10%) received concomitant treatment with fluoxetine during the 6 wk. Three were from the low-dose group and one from the 8-mg group (p = 0.9, Fisher’s exact test, two-tailed).

Prolactin levels

There was a significant increase in prolactin levels from 14.1 ± 9.5 ng/ml at baseline to 51.4 ± 34.8 ng/ml at end-point (t = 3.49, d.f. = 10, p = 0.006) for the 8-mg group, but not for the 2-mg group, where mean prolactin levels only increased marginally from
Table 5. Ratings on Parkinsonism subsection of the ESRS for the two groups over 6 wk (LOCF)

<table>
<thead>
<tr>
<th>Group</th>
<th>2 mg</th>
<th>8 mg</th>
<th>t value</th>
<th>d.f.</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1.9 (±3.1)</td>
<td>0.7 (±1.5)</td>
<td>1.6</td>
<td>38</td>
<td>0.126</td>
</tr>
<tr>
<td>Week 1</td>
<td>5.4 (±11.3)</td>
<td>7.0 (±14.7)</td>
<td>−0.4</td>
<td>37</td>
<td>0.709</td>
</tr>
<tr>
<td>Week 2</td>
<td>4.6 (±10.2)</td>
<td>15.6 (±17.5)</td>
<td>−2.4</td>
<td>37</td>
<td>0.023*</td>
</tr>
<tr>
<td>Week 3</td>
<td>6.0 (±11.1)</td>
<td>15.7 (±16.0)</td>
<td>−2.2</td>
<td>37</td>
<td>0.035*</td>
</tr>
<tr>
<td>Week 4</td>
<td>5.7 (±7.1)</td>
<td>17.9 (±17.2)</td>
<td>−2.9</td>
<td>37</td>
<td>0.007*</td>
</tr>
<tr>
<td>Week 5</td>
<td>5.8 (±7.4)</td>
<td>18.5 (±18.2)</td>
<td>−2.8</td>
<td>37</td>
<td>0.008*</td>
</tr>
<tr>
<td>Week 6</td>
<td>7.3 (±8.7)</td>
<td>18.6 (±18.2)</td>
<td>−2.5</td>
<td>37</td>
<td>0.019*</td>
</tr>
</tbody>
</table>

* Significant at the 0.05 level.

14.1 ±10.1 ng/ml to 15.5 ±8.4 ng/ml over the study period (t = −0.38, d.f. = 9, p = 0.7).

Discussion

This study indicates that acute treatment of first-episode schizophrenia with 2 mg/d haloperidol is as effective as with 8 mg/d. Also, there were no differences in the time to onset of action. This finding supports previous work suggesting efficacy for low doses of haloperidol, and differs from the findings of a previous study in multi-episode patients (Zimbroff et al., 1997) suggesting that efficacy is reduced at lower doses (4 mg/d) compared to higher doses (8 mg/d) of haloperidol.

Both groups of subjects were quite ill at admission, with mean PANSS Total scores in the region of 100 and CGI ratings of >5 (severely ill) at baseline. This is because most patients were recruited from the acute in-patient units at Stikland Psychiatric Hospital, where the more severely ill patients are managed. The fact that 2 mg/d haloperidol was effective in even the most severe cases in our sample suggests that higher doses are not necessary for more severely ill patients.

Our findings provide some evidence to suggest that 2 mg/d haloperidol may be more effective than 8 mg/d in treating negative symptoms, although the relatively small sample limited the power of the study. Thus, while the negative symptom scores in the low-dose group were significantly more improved than the high-dose group at week 2, this was not the case at end-point although there was a 13% difference between the two groups. It has been suggested that first-generation antipsychotics have a very limited treatment effect on negative symptoms and may, in fact, exacerbate these symptoms (Meltzer et al., 1986; Tollefson and Sanger, 1997). This study suggests that the relative lack of efficacy of conventional antipsychotics in treating negative symptoms may be dose-related, i.e. lower doses are better than higher doses. This is consistent with the proposal of a therapeutic window effect for haloperidol (Van Putten et al., 1992), at least with regard to negative symptoms. The poor effect on negative symptoms at high doses may be due to the development of negative symptoms secondary to EPS (Kane and Mayerhoff, 1989; Meltzer et al., 1986), or via some other mechanism. In most previous studies comparing second-generation antipsychotics to haloperidol, the mean dose of haloperidol was considerably higher than 8 mg/d (Tollefson et al., 1997a,b). It may be, therefore, that the reported advantages for the novel antipsychotics in terms of reduction of negative symptoms (Beasley et al., 1996; Breier et al., 1994; Emsley, 1996; Lieberman and the HGDH Study Group, 2000) are spurious.

As anticipated, 2 mg/d haloperidol was better tolerated than 8 mg/d. There were fewer EPS, and prolactin levels were lower in the 2 mg/d group. Also, there were numerically fewer dropouts in this group. Of note however, is that even the 2-mg dose caused a significant increase in parkinsonism scores from baseline. An interesting finding with the prolactin levels was that low-dose treatment had a more profound effect on the prolactin levels of male subjects than on female subjects, whereas the higher dose affected both sexes, but with a greater mean effect in female subjects.

The findings of our study are limited by the small sample size. Also, comparison of low-dose haloperidol with a second-generation antipsychotic would have been of great interest. Nevertheless, these findings provide supportive evidence for the use of low doses of haloperidol rather than high doses, and indicate that this should be the treatment of choice when alternative treatments are not readily available. It is, however, important to note that, while this low-dosing strategy appears to have clear-cut advantages in the short-term treatment of psychosis, recent work of ours (Oosthuizen et al., 2003) indicates that in the longer term, using a low dose of haloperidol does not reduce the risk of tardive dyskinesia. The inability of this strategy to protect against the long-term neurotoxic effects of haloperidol therefore also calls into question earlier proposals by ourselves (Oosthuizen et al., 2001) and others (Geddes et al., 2000) that low doses of first-generation antipsychotics may be a reasonable alternative to the atypical antipsychotics.

Acknowledgements

This study was supported by the Medical Research Council of South Africa. The authors thank Dr Martin Kidd for his statistical advice.
Statement of Interest

None.

Reference


