CHAPTER 3

Efficacy and safety of haloperidol for in-hospital delirium prevention and treatment: a systematic review of current evidence

Edmée J.M. Schrijver, Karola de Graaf, Oscar J. de Vries, Andrea B. Maier, Prabath W.B. Nanayakkara

Abstract

Objective: Haloperidol is generally considered the drug of choice for in-hospital delirium management. We conducted a systematic review to evaluate the evidence for the efficacy and safety of haloperidol for the prevention and treatment of delirium in hospitalised patients.

Methods: PubMed, Embase, Cumulative Index to Nursing and Allied Health (CINAHL), PsycINFO, and the Cochrane Library were systematically searched up to April 21, 2015. We included English full-text randomised controlled trials using haloperidol for the prevention or treatment of delirium in adult hospitalised patients reporting on delirium incidence, duration, or severity as primary outcome. Quality of evidence was graded. Meta-analysis was not conducted because of between-study heterogeneity.

Results: Twelve studies met our inclusion criteria, four prevention and eight treatment trials. Methodological limitations decreased the graded quality of included studies. Results from placebo-controlled prevention studies suggest a haloperidol-induced protective effect for delirium in older patients scheduled for surgery: two studies reported a significant reduction in ICU delirium incidence, one study found a significant reduction in delirium severity and duration. Although placebo-controlled trials are missing, pharmacological treatment of established delirium reduced symptom severity. Haloperidol administration was not associated with treatment-limiting side-effects, but few studies used a systematic approach to identify adverse events.

Conclusion: Although results on haloperidol for delirium management seem promising, current prevention trials lack external validity and treatment trials did not include a placebo arm on top of standard nonpharmacological care. We therefore conclude that the current use of haloperidol for in-hospital delirium is not based on robust and generalisable evidence.
Introduction

Delirium is an acute and fluctuating disturbance in attention, awareness and additionally in cognition (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, DSM-5) [1] occurring as a frequent complication of acute medical illness and hospitalization in elderly patients. Delirium is a common problem on medical and surgical wards. Based on a recent overview of observational studies by Inouye et al.,[2] up to 33% [3] of elderly non-ICU medical patients experience delirium during hospital admission, while reported incidence rates are as high as 51% in elderly hip-fracture patients.[4] The occurrence of delirium is associated with poor patient outcomes and increased health costs.[5] Several studies have demonstrated that delirium predicts worse functional outcomes and institutionalisation for different elderly patient populations.[6-8] Furthermore, development of delirium in elderly inpatients has been linked to post-discharge mortality,[3, 9, 10] for which duration of delirium appears to be an important predictor.[10, 11] Once delirium has established, it is not always reversible, with prolonged and persisting delirium being associated with even poorer outcomes.[12, 13] Therefore, adequate prevention and treatment of delirium is essential.

Management of established delirium primarily includes identifying and treating any underlying cause(s). In addition, nonpharmacological interventions are considered part of standard delirium care, while pharmacological treatments are mostly added to the treatment regimen to reduce the burden of delirium symptoms.[14-16] For prevention of delirium, nonpharmacological interventions are endorsed,[17] yet pharmacological treatments are gaining increased attention even though evidence for their efficacy is limited.[18] Among pharmacological delirium treatments, the typical antipsychotic haloperidol generally is considered first choice in varying patient populations,[14, 16, 19] noting that it has not been approved by the US Food and Drug Administration for this indication.[20]

To date, no systematic review of randomised controlled trials has merely focused first-choice haloperidol for the prevention and treatment of delirium in hospitalised adults. We therefore conducted this systematic literature search to study the efficacy and safety of haloperidol in terms of reducing delirium incidence, duration and/or severity in adult hospitalised patients.

Methods

Data sources

We conducted a systematic literature search up to April 21, 2015 in PubMed, Embase, the Cumulative Index to Nursing and Allied Health (CINAHL), PsycINFO, and
the Cochrane Library, using the index terms and keywords “haloperidol”, “delirium” and “acute confusion”. Identified records were imported into Reference Manager 12 for Microsoft Word® 2003. Duplicate references were removed. The full search strategy is included in Appendix A.

Data collection and quality assessment
The titles and/or abstracts of identified records were independently screened by two reviewers. Studies were considered eligible for inclusion in this review if they: (1) were a randomised controlled trial (RCT); (2) included an intervention group with haloperidol (all routes of administration); (3) included one or more comparison group(s) with either no intervention, placebo, or any other drug (all routes of administration); (3) targeted adult (age 18 years or over) hospitalised patients; (4) targeted incidence, duration, and/or severity of delirium as primary outcome measures. Publications specifically addressing alcohol- or substance-related delirium, patients with schizophrenia, (acute) mania, (psychotic) agitation or aggression were excluded. Only English language full-text articles relevant to the scope of this review and meeting our inclusion criteria were retrieved for detailed evaluation. In case of disagreement between the two reviewers, a third reviewer was consulted. Final article selection was based upon consensus between the investigators. One investigator extracted data on the study design and population, dropout rates, intervention, delirium assessment tools, primary outcome(s), and side-effect reporting. A second investigator checked the acquired data for accuracy and inconsistencies. For each study, p-values and relative risks (RRs) with 95% confidence intervals (CIs) were extracted. If p-values were not reported, these were computed by a statistician using IBM SPSS Statistics 20, p-values less than 0.05 were considered statistically significant. Study characteristics and results were listed by first author, publication year and country. A system based on the Grades of Recommendation, Assessment, Development and Evaluation Working Group (GRADE) and the Cochrane Collaboration’s tool for assessing risk of bias were used to evaluate the quality and applicability of the available evidence (Appendix B).[21, 22] To assess the quality of evidence, each study was assessed independently by two reviewers under the supervision of a third reviewer according to the aforementioned criteria. Because all included studies were RCTs they were initially graded as high quality evidence, and thereafter downgraded based on the presence of risk of bias, resulting in four grade categories: high quality, moderate (downgraded), low (double-downgraded), and very low (triple-downgraded). Final quality rating was based upon consensus between the investigators (Appendix B). We initially planned on conducting a meta-analysis of the included studies, but decided this was not possible due to between-study heterogeneity.
Results

A summary of the search and selection of evidence is provided in Figure 1. A total of 3597 records were identified through our systematic literature search. After duplicate removal, the titles and/or abstracts of 2872 records were screened for eligibility based on our predefined in- and exclusion criteria. We excluded 2855 articles based on a review of the title and/or abstract, mostly because records were not relevant to the scope of this review or not available as English full-text (abstract, other language). As a result, 17 full-text articles were retrieved for detailed evaluation. Two articles were excluded because they did not include a general in-hospital population,[23, 24] one article was excluded because delirium incidence, duration and/or severity were not the primary outcome,[25] and two articles were excluded because they were not randomised,[26, 27] resulting in the inclusion of 12 (four prevention, and eight treatment) studies for this review.

Figure 1. Study flowchart
Study characteristics

Haloperidol for prevention of delirium
We identified four prophylaxis trials. Three studies included a placebo or saline 0.9% control arm, and one study had a no intervention control group. Studies were published between 1999 and 2014, and originated from Japan,[28, 29] The Netherlands,[30] and China.[31] Trials included a total of 1088 (range 80-457) patients all admitted for surgical procedures, predominantly orthopaedic [29, 30] or abdominal [28, 29, 31] surgery. Two studies were initiated in an ICU setting.[28, 31] Two studies specifically excluded patients with profound dementia.[30, 31] Primary outcome for all studies was postoperative delirium incidence assessed with the NEECHAM confusion scale,[29] or DSM criteria for delirium.[28, 30, 31] Other reported outcome measures included delirium duration and severity,[28-30] adverse events,[28-31] time to onset of delirium,[31] hospital length of stay (LOS),[30, 31] ICU LOS,[31] total sleep time,[28] and all-cause 28-day mortality.[31] Two studies were graded as high quality evidence.[30, 31] A summary of the included study characteristics and results are shown in Table 1A.

Haloperidol for treatment of delirium
Our search yielded eight comparison trials, including a total of 486 (range 28-80) patients. Studies were published between 1996 and 2013, and originated from Turkey,[32] USA,[33] India,[34] Korea,[35] Thailand,[36] Japan,[37] Canada,[38] and Greece.[39] Studies compared haloperidol with other typical (chlorpromazine) or atypical (olanzapine, risperidone, quetiapine) antipsychotics, benzodiazepines (lorazepam), tetracyclic antidepressants (mianserin), serotonin 5-HT₃ receptor antagonist (ondasetron), and morphine for the treatment of delirium in hospitalised patients. Three studies enrolled patients who were referred to the hospital consultation-liaison psychiatry service.[34-36] Three studies explicitly excluded patients with profound cognitive impairment or dementia.[32, 34, 35] One study did not use a valid tool for delirium diagnosis.[39] Two trials excluded patients who were diagnosed with hypoactive delirium.[32, 36] Primary efficacy outcome was delirium severity assessed with the DI,[38] DRS,[33, 37] DRS-R-98,[34, 36] MDAS,[35] and RASS respectively.[32] Other outcomes included adverse events,[32-38] time to onset of delirium,[32] time to response,[35, 36] delirium duration,[32] total sleep time,[36] hospital LOS,[32] ICU LOS,[32] and mortality.[32, 33, 36] One study was graded as high quality evidence.[36] Study details are provided in Table 1B.
### Table 1a. Study characteristics and results of prevention RCTs included in this review

<table>
<thead>
<tr>
<th>Author, year (country)</th>
<th>Study quality</th>
<th>Study design</th>
<th>Study population (n)</th>
<th>Dropouts (n)</th>
<th>Age mean±SD (y)</th>
<th>Intervention</th>
<th>Treatment of established delirium</th>
<th>Delirium assessment</th>
<th>Efficacy outcome</th>
<th>Safety assessment</th>
<th>Safety outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fukata, 2014 (Japan)</td>
<td>Low</td>
<td>Multi-centre, open-label RCT</td>
<td>Elective abdominal (&gt;80%) and orthopedic surgery patients (121), 52.9% male, mean MMSE HAL 23.3±0.7 and CON 23.0±0.7, included between January 2007 and December 2012</td>
<td>2 (1.7%)</td>
<td>HAL 80.2±0.5 NO-I: 80.5±0.5</td>
<td>HAL 2.5mg opd iv for 3 days vs. NO-I initiated on post-operative day 1</td>
<td>Conventional treatments at flexible dosages (mostly HAL)</td>
<td>NEECHAM score (cutoff score &lt;20) daily until 7th postoperative day</td>
<td>Incidence: HAL = NO-I (42.4 vs. 33.3%, 95% CIs 29.6-55.9 vs. 21.7-46.7%; p=0.309) Duration: HAL = NO-I (3.4 vs. 1.1, 95% CIs 0.8-1.95 vs. 0.6-1.6 days; p=0.356) Severity: HAL = NO-I (-)</td>
<td>-</td>
<td>No intervention-related adverse events noted</td>
</tr>
</tbody>
</table>

<p>| Kalisvaart, 2005 (Netherlands) | High          | Single-centre, double-blind, PLA-controlled, RCT | Acute and elective (73.5%) hip-surgery patients at-risk for delirium on surgical ward (430), 20.2% male, mean MMSE HAL 25.0±6.9 and PLA 24.5±4.2, included between | 48 (11.2%) | HAL 78.7±6.0 PLA 79.6±6.3 | HAL 0.5mg po tid for max 6 days vs. PLA (initiated on admission, max delay for surgery 72h) | HAL and/or LOR tid at flexible dosages | DSM-IV and CAM daily, DRS-R-98 for severity | Incidence: HAL = PLA (RR = 0.91, 95%CI = 0.59-1.44; p&lt;0.687*). Duration: HAL &lt; PLA (mean difference 6.4, 95%CI = 4.0-8.0 days; p&lt;0.001) | BAS, ECG | No EPS (BAS 0 for all patients), no drug-related side effects noted |</p>
<table>
<thead>
<tr>
<th>Author, year (country)</th>
<th>Study quality</th>
<th>Study design</th>
<th>Study population (n)</th>
<th>Dropout (n)</th>
<th>Age mean±SD (y)</th>
<th>Intervention</th>
<th>Treatment of established delirium</th>
<th>Delirium assessment</th>
<th>Efficacy outcome</th>
<th>Safety assessment</th>
<th>Safety outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaneko, 1999 (Japan)</td>
<td>Low</td>
<td>Single-centre, not-blinded, RCT</td>
<td>Elective gastrointestinal surgery patients (80), 64.1% male, CI HAL 5.3% and SAL 10.0%, included between April 1995 and August 1998</td>
<td>2 (2.5%)</td>
<td>HAL 72.4±8.2 SAL 73.1±9.3</td>
<td>HAL 5mg iv for 5 days vs. SAL initiated on postoperative day 1</td>
<td>HAL at flexible dosages</td>
<td>DSM-III-R on 5th postoperative day</td>
<td>Incidence: HAL &lt; SAL (10.5% vs. 32.5%; p=0.019) Duration: - Severity: -</td>
<td>-</td>
<td>No intervention-related adverse events noted</td>
</tr>
<tr>
<td>Wang, 2012 (China)</td>
<td>High</td>
<td>Two-centre, double-blind, PLA-controlled, RCT</td>
<td>Non-cardiac surgery patients in ICU (457), male 63.0%, included between June 2009 and May 2010</td>
<td>0</td>
<td>HAL 74.0±5.8 SAL 74.4±7.0</td>
<td>HAL 0.5mg iv bolus and 0.1mg/h iv for 12h vs. SAL (PLA) initiated postoperatively &lt; 1h after enrollment</td>
<td>Open-label HAL (initial dose 0.5-1 mg iv) for severe agitation</td>
<td>RASS score &gt; -4, CAM-ICU daily until 7th postoperative day</td>
<td>Incidence: HAL &lt; SAL (15.3% vs. 23.2%; p=0.031)</td>
<td>-</td>
<td>ECG No EPS, QT prolongation in both groups (HAL 1.7%, PLA 2.2%; p = 0.995)</td>
</tr>
</tbody>
</table>
### Table 1b. Study characteristics and results of treatment RCTs included in this review

<table>
<thead>
<tr>
<th>Study</th>
<th>Study quality</th>
<th>Study design</th>
<th>Study population (n)</th>
<th>Drop-outs (n)</th>
<th>Age mean±SD (y)</th>
<th>Intervention</th>
<th>Rescue medication</th>
<th>Delirium assessment</th>
<th>Efficacy outcome</th>
<th>Safety assessment tools</th>
<th>Safety outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atalan, 2013 (Turkey)</td>
<td>Very low</td>
<td>Single-centre, not-blinded RCT</td>
<td>Cardiac surgery patients with hyperactive delirium (53), 73.6% male, RASS HAL 2.9 and MOR 3.0, included between January 2010 and July 2012</td>
<td>0</td>
<td>HAL 66±8.4 MOR 65.7±9.7</td>
<td>HAL vs. MOR sulfate 5mg im q1h (max 20mg/d)</td>
<td>LOR 2.5mg bid po</td>
<td>DSM-IV, CAM-ICU, RASS (cutoff score &gt;2) every 12h until discharge or 10th postoperative day</td>
<td>Severity: HAL = MOR (RASS 0.04 vs. 0.0; p=0.327) Duration: HAL = MOR (33.9±16.7 vs. 31.6±16.6 hrs; p=0.607)</td>
<td>-</td>
<td>No between-group differences in serious adverse effects. Mortality: HAL 7.7% vs MOR 3.7%; p=0.610</td>
</tr>
<tr>
<td>Breitbart, 1996 (USA)</td>
<td>Moderate</td>
<td>Single-centre, double-blind, 6-day RCT</td>
<td>AIDS patients with delirium (30), 77% male, mean MMSE 12.7±7.6, mean DRS 20.1±3.5</td>
<td>0</td>
<td>39.2±8.8</td>
<td>HAL vs. CHL vs. LOR at flexible dosages for up to 6 days</td>
<td>-</td>
<td>DSM-III-R, DRS (cutoff score ≥13) daily</td>
<td>Severity: HAL = CHL = LOR (changes in DRS scores: main effect for time p&lt;0.001; for drug p&lt;0.44; drug-by-time interaction p&lt;0.07)</td>
<td>ESRS, SESC</td>
<td>Treatment-limiting adverse effects in LOR group</td>
</tr>
<tr>
<td>Grover, 2011 (India)</td>
<td>Very low</td>
<td>Single-centre, single-blind RCT</td>
<td>Patients referred to consultation-liaison psychiatry service (74), 70.3% male, mean DRS-R-98 HAL 21.9±4.8, OLA 22.6±4.5 and RIS 23.8±5.2</td>
<td>10 (13.5%)</td>
<td>HAL 44.1±16.8 OLA 45.4±19.2 RIS 46.5±14.5</td>
<td>HAL 0.25-10mg vs. OLA 0.25-4mg vs. RIS 1.25-20mg</td>
<td>Same drug for HAL and OLA group; HAL or LOR for RIS group, iv at flexible dosages</td>
<td>CAM, DRS-R-98 daily for 6 days</td>
<td>Severity: HAL = OLA = RIS (mean DRS-R-98 scores: between-group difference at day 6 p=0.424)</td>
<td>SAS, AIMS, UKU</td>
<td>Occurrence of side-effects: HAL 19%, OLA 8.7%, and RIS 30% (p&gt;0.05)</td>
</tr>
<tr>
<td>Han, 2004 (Korea)</td>
<td>Moderate</td>
<td>Single-centre, double-blind RCT</td>
<td>Patients referred to consultation-liaison psychiatry service (28), 54.2% male, mean DRS HAL</td>
<td>4 (14.3%)</td>
<td>HAL 66±15.9 RIS 65.6±8.3</td>
<td>HAL (initial dose 0.75 mg bid) vs. RIS (initial dose 0.5)</td>
<td>-</td>
<td>DSM-III-R (SCID), CAM, DRS (cutoff score ≥13), MDAS daily</td>
<td>Severity: HAL = RIS (mean MDAS: main effect for time p&lt;0.05)</td>
<td>-</td>
<td>HAL: mild akathisia (n=1). No clinically significant</td>
</tr>
<tr>
<td>Study</td>
<td>Study quality</td>
<td>Study design</td>
<td>Study population (n)</td>
<td>Drop-out (n)</td>
<td>Age mean±SD (y)</td>
<td>Intervention</td>
<td>Rescue medication</td>
<td>Delirium assessment</td>
<td>Efficacy outcome</td>
<td>Safety assessment tools</td>
<td>Safety outcome</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------------</td>
<td>-----------------------------------</td>
<td>----------------------</td>
<td>--------------</td>
<td>-----------------</td>
<td>--------------</td>
<td>-------------------</td>
<td>---------------------</td>
<td>-----------------</td>
<td>-----------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Maneeton, 2013 (Thailand)</td>
<td>High</td>
<td>Single-centre, double-blind, 7-day RCT</td>
<td>Patients referred to consultation-liaison psychiatry service (52), 67.3% male, mean DRS-R-98 HAL 29.7±4.6 and QUE 29.0±4.4, included between June 2009 and April 2011</td>
<td>b</td>
<td>HAL 57.0±11.9 QUE 56.6±12.0</td>
<td>HAL 0.5-2.0 mg/day vs. QUE 25-100 mg/day po initiated 3.1±2.7 days after delirium diagnosis</td>
<td>-</td>
<td>DSM-IV-TR, CAM, DRS-R-98 daily for 7 days</td>
<td>between-group difference p=0.51; group-by-time effect p=0.14,</td>
<td>Safety: HAL = QUE (between-group difference in mean DRS-R-98 scores; p=0.39)</td>
<td>MSAS: Mean MSAS scores: HAL 0.3±1.1 and QUE 0.3±0.7 (p=0.51). Mortality: n=1 per group, not intervention-related</td>
</tr>
<tr>
<td>Nakamura, 1997 (Japan)</td>
<td>Very low</td>
<td>Single-centre, open-label RCT</td>
<td>Hospitalised patients, 57.6% postoperative, (66), 71.2% male, mean DRS HAL 22.1±3.8 MIA 21.±4.1</td>
<td>-</td>
<td>HAL 67.7±15.0 MIA 63.8±12.8</td>
<td>HAL 2-6 mg/day vs. MIA 10-60 mg/day po</td>
<td>-</td>
<td>DSM-IV, DRS (cut-off score ≥16) daily</td>
<td>Severity: HAL = MIA (between-group difference in DRS scores; -)</td>
<td>No formal instruments used</td>
<td>MIA: oversedation (n=2), no other side-effects noted</td>
</tr>
<tr>
<td>Skrobik, 2004 (Canada)</td>
<td>Very low</td>
<td>Single-centre, single-blind, 5-day RCT</td>
<td>Medical (55.5%) and surgical (65.8%) elective ICU patients (80), 72.6% male, included between July 2000 and September 2001</td>
<td>7 (8.8%)</td>
<td>HAL 63.±11.7 OLA 67.5±6.°</td>
<td>HAL (initial dose 2.5-5 mg q8h) vs. OLA (initial dose 5 mg/d), dosage adjusted for age ≥60, initiated &lt; 2h of delirium diagnosis</td>
<td>HAL IV and BEN at flexible dosages</td>
<td>DSM-IV, ICU-DSC (cutoff score ≥4), DI Severity: HAL = OLA (between-group difference in DI scores, interaction effect; p=0.64)</td>
<td>Ross-Chouinard, Angus-Simpson scale</td>
<td>HAL: minor EPS (n=6)</td>
<td></td>
</tr>
<tr>
<td>Tagarakis, 2012 (Greece)</td>
<td>Very low</td>
<td>Single-centre, RCT</td>
<td>On-pump cardiac surgery patients (80), 66.3% male, included between 2008 and 2010</td>
<td>-</td>
<td>HAL 70.9±9.9 OND 70.1±9.3</td>
<td>HAL 5 mg vs. OND 8mg iv sd</td>
<td>-</td>
<td>No formal instrument (4-point scale), once 10min after intervention</td>
<td>Severity: HAL = OND (between-group difference in 4-point scale improvement; -)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 1b. (Continued)

RCT: Randomised Controlled Trial; HAL: haloperidol; NO-I: no intervention; PLA: placebo; LOR: lorazepam; SAL: saline 0.9%; MOR: morphine; CHL: chlorpromazine; OLA: olanzapine; RIS: risperidone; QUE: quetiapine; MIA: mianserin; BEN: benzodiazepines; OND: ondasetron; MMSE: Mini-Mental State Examination; NEECHAM: Neelon/Champagne Confusion Scale; RASS: Richmond Agitation Sedation Scale; CAM: Confusion Assessment Scale; DRS(-R-98): Delirium Rating Scale(-Revised-98); DI: Delirium Index; ICU-DSC: ICU Delirium Screening Checklist; SCID: Structured Clinical Interview for DMS-III-R; ESRS: Extrapyramidal Symptom Rating Scale; SESC: Side Effects and Symptoms Checklist; (M)SAS: (Modified) Simpson Angus Scale; AIMS: Abnormal Involuntary Movement Scale; UKU: Udvalg for Kliniske Undersogelser side effect rating scale; CGI-S: Clinical Global Impression-Severity; CI: cognitive impairment; ICU: Intensive Care Unit; mg: milligrams; opd: once a day; bid: twice a day; tid: three times a day; q*h: every * hour(s); min: minute(s); sd: single dose; po: oral; iv: intravenous; im: intramuscular; h: hour(s); vs: versus; BAS: Barnes Akathisia Scale; ECG: electrocardiogram; -: not reported.

a Intention-to-treat analysis performed on all 430 randomised patients, 35 patients (8.1%) were lost to follow-up for primary outcome assessment; p-value (for chi-square test) computed by authors.
b Intention-to-treat analysis performed on all 52 randomised patient, 17 patients (32.7%) did not complete the study.
c 73 patients included in final analysis, 61 patients alive and in-hospital for >3 days, accounting for difference between randomised and 3-day analysis.
* Significant between-group difference in baseline characteristic (p=0.046).

Study results on the primary outcome measures

Delirium incidence
Postoperative delirium incidence ranged from 15.8% [30] to 37.8% [29] across prevention studies. Data from these studies indicated that haloperidol prophylaxis significantly reduced postoperative delirium incidence in patients admitted to the ICU predominantly after abdominal surgery.[28, 31] No significant effect was demonstrated after mostly elective abdominal and orthopedic surgery in older, at-risk patients.[29, 30]

Delirium duration
Three prevention studies reported on delirium duration.[28-30] Two of these studies found a significant reduction in the duration of postoperative delirium for patients that received haloperidol prophylaxis when compared with placebo.[28, 30] One study did not report any values.[28] Kalisvaart et al. found a shorter delirium
duration of 6.4 days in (mostly) elective hip-surgery patients who had received perioperative haloperidol prophylaxis (5.4±4.9 versus 11.8±7.6 days for haloperidol and placebo respectively, 95% CI 4.0-8.0; p < 0.001).[30] Overall, treatment studies demonstrated that delirium symptoms resolved within seven days of initiation of therapy, irrespective of the drug treatment received.[32-39]

**Delirium severity**
Delirium severity was assessed with different scales, and studies varied in baseline delirium severity scores. Three prevention studies reported on delirium severity.[28-30] Two of these studies found that symptom severity during a delirium episode was significantly lower in patients who had received haloperidol prophylaxis compared to placebo, one of them did not report any values.[28, 30] All of the included treatment studies demonstrated that delirium symptoms significantly improved after drug therapy, but that the decrease in delirium severity scores from baseline did not significantly differ between the treatment groups. In addition, in studies reporting on the average time to response, remission and response rates at the end of the study period were not significantly different between groups, although different definitions and cut-off points were used. Average periods before response varied between 1.9 and 4.2 days for haloperidol, which was comparable with second-generation antipsychotics.[35, 36] Breitbart et al. demonstrated that delirium symptoms improved to values below the diagnostic threshold for delirium (DRS ≤13) within the first 24 hours after treatment initiation for both neuroleptic drug groups.[33] Response rates at the end of the study period ranged from 70.6% to 78.6% for haloperidol, and from 42.0% to 79.2% for comparison groups.[35-37] Remission rates varied between 67.9% and 100% for haloperidol, and for comparison groups between 69.6% and 100%.[32, 34, 36]

**Haloperidol safety**
Six of the 12 studies that we included in this review used standardized methods to record side-effects, mainly focusing on EPS. Details on safety assessment and outcomes for all included studies are listed in Table 1A and B. Haloperidol use was generally safe, since none of the studies reported any serious intervention-related or treatment-limiting side-effects during the study period. Two placebo-controlled prevention trials described standard QTc interval monitoring.[30, 31, 36] QTc prolongation (>60ms, or QTc interval >500ms) was recorded in 1.9% (9/457) of older postoperative ICU patients after a total daily iv haloperidol dose 1.7 mg, with no significant difference compared with placebo (p = 0.995).[31] No QTc prolongation was observed in older patients who underwent hip-surgery.[30] All-cause mortality rates were not higher for patients receiving haloperidol compared with placebo or other drugs.[31-33, 36]
Our review demonstrates that since the early 1990s, 12 randomised trials have been conducted. The results of these studies do not demonstrate an apparent superior efficacy or safety of haloperidol for either the prevention or treatment of hospital-associated delirium in adult patients as compared with placebo or other drugs. We identified several methodological limitations in the studies included for this review that evaluated the efficacy of haloperidol both for the prevention and treatment of hospital-associated delirium. Most importantly, while all prevention studies included either a no-intervention or placebo arm, our search yielded no placebo-controlled studies for delirium treatment. We have highlighted these limitations in Appendix B.

The primary aim of delirium prevention studies was to reduce postoperative delirium incidence. Data from the placebo-controlled prevention studies included for this review indicate that haloperidol prophylaxis may be effective in reducing postoperative delirium only in older patients admitted to an ICU after surgery.[28, 31] However, we graded one of the two studies who demonstrated these results as low quality as this study was not blinded, randomization was not adequately described other than that a closed envelope system was used, the sample size was small, and two patients were excluded from analysis for reasons that were not clear. [28] Wang et al. performed a placebo-controlled RCT that we graded as high quality, although strict trial eligibility criteria limit the external validity or generalisability of their results as they only included postoperative patients aged ≥65 years who were admitted to an ICU after mostly intra-abdominal surgery.[31]

To evaluate the efficacy of haloperidol for in-hospital delirium management, reducing the duration and severity of delirium were, next to delirium incidence, also considered important outcome measures for this review. The data arising from the included prevention studies showed contradicting results on the effect of haloperidol prophylaxis on postoperative delirium duration and severity. Kalisvaart et al.[30] showed that perioperative haloperidol prophylaxis on top of nonpharmacological multicomponent interventions reduced delirium duration and severity in older at-risk patients scheduled for (mostly elective) hip-surgery, while Fukata et al.[29] found contradicting results for older patients after elective abdominal and orthopedic surgery. Differences in study design, delirium assessment tools and methods used to classify at-risk patients may have contributed to the differences in study outcomes. Although the positive findings in the study conducted by Kalisvaart et al.[30] were statistically significant, one might question whether they are also clinically relevant. The highest DRS-R-98 scores observed during a delirium episode in the study by Kalisvaart et al.[30] were relatively low (18.4 and 14.4 for haloperidol and placebo, respectively) compared with average
scores in delirious subjects enrolled in two validation studies of the DRS-R-98 score (approximately 21.3 and 22).[40, 41] The observed mean 4-point reduction in the maximum delirium severity score (DRS-Max) in the perioperative haloperidol group may therefore be of lesser clinical relevance. Furthermore, Kalisvaart et al.[30] demonstrated a significant reduction in delirium duration in patients who had received perioperative haloperidol prophylaxis compared with placebo (mean difference of 6 days, 5.4 days for haloperidol versus 11.8 days for placebo). However, mean DRS-R-98 scores were only significantly lower in the perioperative haloperidol group during the first three days after delirium onset, compared with placebo. As of the 4th day, DRS-R-98 scores seemed to reach a plateau phase in both groups with eyeball-estimate mean scores of 14, suggesting that from that point the clinical course of delirium was comparable across groups. These results indicate that the additional days spent in delirium for patients who had received perioperative placebo were not characterised by more severe delirium symptoms as assessed with the DRS-R-98. The limited generalisability and contradicting results of these findings support the need for future well-designed placebo-controlled trials to investigate the use of prophylactic haloperidol on top of nonpharmacological interventions in delirium prevention in a general hospital population.

Results from the treatment studies identified through our systematic literature search suggest that improvement in delirium symptoms occurs irrespective of the drug administered, with no superior efficacy of haloperidol over second-generation antipsychotics or other drugs, bearing in mind that included comparison studies are too few and heterogeneous to draw valid conclusions.[32-39] Prospective non-randomised and retrospective studies excluded from this review show similar results.[27, 42-45] Data from two double-blind placebo-controlled RCTs, which we excluded from this review because the primary outcome was number of days alive without delirium or coma, also support the general conclusions of our review.[25, 46] Both trials included medical and surgical ICU patients, irrespective of delirium or coma status at baseline, limiting the ability to define these studies as either prevention or treatment. Girard et al. randomised 103 ICU patients to receive 5mg haloperidol, 40mg ziprasidone, or matching placebo per os at flexible dosages.[46] No between-groups differences were found on primary and secondary outcomes, including delirium days (median 4, interquartile range (IQR) 2-7, 2-8, and 2-6 days for haloperidol, ziprasidone and placebo group respectively; p = 0.93), and 21-day mortality (11%, 13% and 17% for haloperidol, ziprasidone, and placebo group respectively; p = 0.81).[46] Page et al. enrolled 142 ICU patients ≤72 hours of admission and administered haloperidol 2.5mg or placebo (saline 0.9%) iv every 8 hours.[25] The number of days spent in delirium during the 14 day study period was not different between groups (median 5, IQR 2-8 and 1-8 days for haloperidol and placebo group respectively; difference 0.01, 95% CI -1.3-1.3; p = 0.99), neither
was 28-day mortality (RR 1.04, 95% CI 0.6-1.8).[47] Also, EPS (e.g. akathisia) and QTc prolongation occurrence rates did not differ significantly between haloperidol and placebo groups.[25, 46] Overall, in absence of placebo or non-pharmacological intervention control groups, we conclude that available evidence from included randomised treatment trials is insufficient to support the current extensive use of haloperidol for in-hospital delirium treatment.

In addition to the outcomes that were of primary interest for this review (i.e. delirium incidence, duration and/or severity), included studies also reported on other clinically relevant outcomes. First, a disrupted sleep-wake cycle is common in hospitalised patients who develop delirium. Included studies who reported on this outcome indeed found significant differences in day and night sleeping times between delirious and non-delirious patients, but prophylactic haloperidol did not increase total average sleep time compared with placebo or quetiapine.[28, 36] Total sleep time was inversely correlated with DRS-R-98 severity scores, and increased during treatment irrespective of the drug received.[36] The limited amount of evidence from the studies we included for this review support the presence of a disrupted sleep-wake cycle during a delirium episode, and that it is likely to be restored when delirium resolves, independent of the chosen drug treatment. Second, development of delirium during hospital stay has been associated with negative outcomes such as increased LOS.[47] Kalisvaart et al. found that for patients who experienced delirium after hip-surgery, the number of hospital days was on average 5.5 days less if they had received early low-dose haloperidol prophylaxis compared with placebo (17.1±11.1 versus 22.6±16.7 days for haloperidol and placebo respectively; p < 0.001).[30] However, mean hospital days in the total intention-to-treat group were not significantly different (13.8±7.7 versus 13.6±7.8 for haloperidol and placebo, respectively; p = 0.84). For this study they recruited participants, of which only approximately 25% were acute admissions, between August 2000 and 2002. Since then, much has changed in the care for acute geriatric hip-surgery patients. A majority of these patients are operated within 48 hours of admission,[48] and average hospital LOS for older patients after hip-surgery has drastically decreased to approximately four to seven days.[48, 49] These changes indicate that the results of the study conducted by Kalisvaart et al. may not apply to the current geriatric hip-surgery population. Hence in our opinion, a future placebo-controlled trial could reveal results that may differ from those of the trial conducted by Kalisvaart et al. For patients admitted to the ICU after non-cardiac surgery, postoperative haloperidol prophylaxis did not reduce median hospital LOS in the total study population, nor in a subpopulation of patients who did develop delirium.[31] However, postoperative haloperidol prophylaxis did reduce postoperative ICU LOS (21.3, 95% CI 20.3-22.2 versus placebo 23.0, 95% CI 20.9-25.1 hours; p = 0.024).[31] One could debate on whether this difference of approximately two hours in ICU LOS is clinically relevant
considering the fact that the majority of these patients were discharged from the ICU within 24 hours, although the difference was more evident when limited to the subpopulation of patients who developed delirium (haloperidol 19.6, 95% CI 16.3-22.9 versus placebo 41.4, 95% CI 39.3-43.5 hours; p = 0.006).[31] Treatment of ICU delirium with haloperidol when compared with morphine did not reduce mean ICU LOS (3.3±2.3 versus 2.9±1.5 days; p = 0.402) or hospital LOS (8.5±3.4 versus 8.9±3.1 days; p = 0.607).[32] Overall, studies report contradictory results on the effect of haloperidol prophylaxis on several important clinical outcomes, and study population characteristics limit the generalizability of results to a wider in-hospital population. These trial findings do indicate that in selected patient populations early prophylactic haloperidol may accelerate resolution of established delirium, and in our opinion support the argument in favor of future placebo-controlled trials in a more general in-hospital patient population.

Safety aspects of haloperidol use were systematically evaluated in six of the 12 studies included in this review. Administration of haloperidol is known to carry a risk of EPS and QTc prolongation with a risk of sudden cardiac death.[50, 51] Neither preventative nor treatment studies found a significant increase in haloperidol-related adverse events or mortality, compared to placebo or other drugs. Because not all studies used standardised adverse event monitoring, this may have resulted in underreporting of events. The two prevention trials reporting on standardised ECG monitoring reported no life threatening QTc prolongation or torsade de pointes. [30, 31] To our knowledge, there is no consensus on whether routine ECG recording is necessary prior to or during haloperidol administration. Future well-designed studies should include protocolized, repeated (cardiac) adverse event monitoring to extend existing knowledge base.

This systematic literature review is limited by potential publication bias, since only English language full-text RCTs were selected, and the possibility of unpublished negative data. With our broad search strategy and predefined criteria we aimed to provide an up-to-date review of all available and relevant RCTs on the efficacy and safety of haloperidol for hospital-associated delirium in adult patients.

**Conclusions**

In conclusion, haloperidol is considered the drug of choice for hospital-associated delirium by many healthcare professionals and professional guidelines.[14, 15, 52, 53] The data presented in this systematic review do not support the current extensive use of haloperidol for this indication. Although results from placebo-controlled prevention studies suggest that haloperidol prophylaxis may prevent postoperative ICU delirium in older patients, these study results may not apply to
other populations. None of the selected treatment trials included a no-intervention or placebo arm to evaluate the natural course and resolution of delirium symptoms without pharmacological treatment on top of standard nonpharmacological care. Haloperidol use was generally proven to be safe even in frail older patients, although few studies used standardized methods to monitor side-effects (EPS and QTc prolongation in particular).

In our opinion, result from this review support the need for further well-designed, placebo-controlled clinical trials to evaluate efficacy and safety of haloperidol on top of standard nonpharmacological interventions for the prevention and treatment of delirium in a general (at-risk) in-hospital patient population.
Appendix A

Search strategy used for the literature review. These databases were searched up to April 21, 2015.

<table>
<thead>
<tr>
<th>Database</th>
<th>Search #1</th>
<th>Search #2</th>
<th>Search #3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Embase</td>
<td>‘haloperidol’/exp OR ‘haloperidol’ OR haloperidol:ab,ti OR haldol:ab,ti</td>
<td>‘delirium’/exp OR delir*:ab,ti OR delier:ab,ti OR (acute NEXT/1 confusion*):ab,ti</td>
<td>#1 AND #2</td>
</tr>
<tr>
<td>CINAHL</td>
<td>(MH “Haloperidol”) OR TI (haloperidol OR haldol) OR AB (haloperidol OR haldol)</td>
<td>(MH “Delirium” OR TI (delir* OR delier OR “acute confusion*”) OR AB (delir* OR delier OR “acute confusion*”))</td>
<td>#1 AND #2</td>
</tr>
<tr>
<td>PsycINFO</td>
<td>DE “Haloperidol” OR TI (haloperidol OR haldol) OR AB (haloperidol OR haldol)</td>
<td>(DE “Delirium” OR TI (delir* OR delier OR “acute confusion*”) OR AB (delir* OR delier OR “acute confusion*”))</td>
<td>#1 AND #2</td>
</tr>
<tr>
<td>Cochrane Library</td>
<td>DE “Haloperidol” OR TI (haloperidol OR haldol) OR AB (haloperidol OR haldol)</td>
<td>DE “Haloperidol” OR TI (haloperidol OR haldol) OR AB (haloperidol OR haldol)</td>
<td>#1 and #2</td>
</tr>
</tbody>
</table>
## Appendix B

Quality assessment of RCTs included for this review based on the Grades of Recommendation, Assessment, Development and Evaluation Working Group (GRADE) and the Cochrane Collaboration’s tool for assessing risk of bias.[21, 22]

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Placebo-controlled</th>
<th>Random sequence generation</th>
<th>Group similarity</th>
<th>Allocation concealment</th>
<th>Blinding of patients and care providers</th>
<th>Blinding of outcome assessor</th>
<th>Incomplete outcome data</th>
<th>Selective outcome reporting</th>
<th>Overall study quality</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prevention</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fukata, 2014</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>No</td>
<td>Low</td>
</tr>
<tr>
<td>Kalisvaart, 2005</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>High</td>
</tr>
<tr>
<td>Kaneko, 1999</td>
<td>Yes</td>
<td>Yes^b</td>
<td>Yes</td>
<td>Yes^b</td>
<td>No^b</td>
<td>No^b</td>
<td>Yes^b</td>
<td>Yes^b</td>
<td>Low</td>
</tr>
<tr>
<td>Wang, 2012</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>High</td>
</tr>
<tr>
<td><strong>Treatment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atalan, 2013</td>
<td>No</td>
<td>Unclear^c</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Unclear^c</td>
<td>No</td>
<td>No</td>
<td>Very low</td>
</tr>
<tr>
<td>Breitbart, 1996</td>
<td>No</td>
<td>Unclear^d</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Unclear^d</td>
<td>No</td>
<td>No</td>
<td>Moderate</td>
</tr>
<tr>
<td>Grover, 2011</td>
<td>No</td>
<td>Yes</td>
<td>No^a</td>
<td>Unclear^e</td>
<td>No^a</td>
<td>Yes</td>
<td>Unclear^e</td>
<td>No</td>
<td>Very low</td>
</tr>
<tr>
<td>Han, 2004</td>
<td>No</td>
<td>Unclear^f</td>
<td>Yes</td>
<td>Yes^f</td>
<td>Yes^f</td>
<td>Yes^f</td>
<td>No</td>
<td>No</td>
<td>Moderate</td>
</tr>
<tr>
<td>Maneeton, 2013</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>High</td>
</tr>
<tr>
<td>Nakamura, 1997</td>
<td>No</td>
<td>Unclear^g</td>
<td>Unclear^g</td>
<td>No^g</td>
<td>No</td>
<td>No</td>
<td>Yes^g</td>
<td>No</td>
<td>Very low</td>
</tr>
<tr>
<td>Skrobik, 2004</td>
<td>No</td>
<td>No^h</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes^h</td>
<td>No</td>
<td>No</td>
<td>Very low</td>
</tr>
<tr>
<td>Tagarakis, 2012</td>
<td>No</td>
<td>Unclear^i</td>
<td>Yes</td>
<td>No^i</td>
<td>No^i</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Very low</td>
</tr>
</tbody>
</table>
Appendix B. (Continued)
Random sequence generation: “Yes” indicates random component in sequence generation procedure described in sufficient detail (e.g. computer random number generation, envelopes), “No” indicating a non-random component (e.g. admission day), or “Unclear” indicating insufficient information on randomization methods; Group similarity at baseline: “Yes” indicating similarity at baseline between groups for important patient characteristics and demographics, or “No”; Allocation concealment: “Yes” indicating adequate allocation concealment prior to assignment to the intervention (e.g. central allocation including web-based, pharmacy-controlled, sealed envelopes), “No”, or “Unclear” indicating insufficient detail on methods used to conceal allocation sequence; Blinding of patients and care providers: “Yes” indicating double-blinding was ensured and described in sufficient detail, or “No” indicating single-blind, open-label study; Blinding of outcome assessment: “Yes” indicating outcome assessor(s) unaware of allocated intervention(s), “No” indicating not-blinded, or “Unclear”; Incomplete outcome data: “Yes” indicating number and/or reason for withdrawal/dropout not described, not comparable between groups, or more than 20%, “No” indicating all subjects were accounted for in the analysis (ITT analysis) with no more than 20% drop-outs, number and reasons for withdrawal/dropout were described and comparable between groups, or “Unclear” indicating no ITT analysis, but number and reasons for withdrawal/dropouts were described and ≤20%; Selective outcome reporting (reporting bias): “No” indicating pre-specified primary and secondary outcomes, “Yes” indicating one or more primary outcomes were not pre-specified, outcomes of interest were reported incompletely, and/or no results for expected key outcome.

ITT: Intention-To-Treat.

a Intention-to-treat analysis performed; per-protocol analyses included 382 patients although from the dropouts (48, haloperidol n=20, placebo n=28), 13 (haloperidol n=9, placebo n=4) patients did complete follow-up for primary outcome assessment b Randomization by way of a closed envelope system, unclear whether envelopes were sequentially numbered; blinding not reported; missing data and exclusion for analysis of two patients in haloperidol group, abstruseness on whether this occurred before (as mentioned in the text) or after (as reported in the table) randomization; selective reporting for outcome measures delirium intensity and duration (no severity scale and/or score reported), only reported as ‘not significant’.

c Insufficient information on sequence generation process, description only includes ‘randomised’; clinical evaluation was made by intensivist, together with psychiatrist who was blinded to the study groups, suggesting there was no complete blinding of outcome assessment.
References


