Review

Run-in periods and clinical outcomes of antipsychotics in dementia: A meta-epidemiological study of placebo-controlled trials

Tessa A. Hulshof1 | Sytse U. Zuidema1 | Christine C. Gispen-de Wied2 | Hendrika J. Luijendijk1

1Department of General Practice and Elderly Care Medicine, University Medical Center Groningen, Groningen, The Netherlands
2Dutch Medicines Evaluation Board, Utrecht, The Netherlands

Correspondence
H. J. Luijendijk, Department of General Practice and Elderly Care Medicine, University Medical Center Groningen, PO Box 196, 9700 AD Groningen, The Netherlands.
Email: h.j.luijendijk@umcg.nl

Funding information
Dutch Medicines Evaluation Board

Abstract

Purpose: Run-in periods are used to identify placebo-responders and washout. Our aim was to assess the association of run-in periods with clinical outcomes of antipsychotics in dementia.

Methods: We searched randomized placebo-controlled trials of conventional and atypical antipsychotics for neuropsychiatric symptoms (NPS) in dementia in electronic sources and references of selected articles. We extracted (a) the presence of a run-in period, use of placebo/investigated drug during run-in (versus washout only), and run-in duration (1 week or more) and (b) the reduction in NPS, number of participants with somnolence, extrapyramidal symptoms (EPS), and deaths per treatment group. We pooled clinical outcomes comparing antipsychotic and placebo groups in trials with and without run-in.

Results: We identified 35 trials. Twenty-nine trials used run-in. The pooled standardized mean difference in the reduction of NPS was −0.170 (95% CI, −0.227 to −0.112) in trials with run-in and −0.142 (95% CI, −0.331 to 0.047) in trials without run-in. The pooled odds ratio for somnolence was 2.8 (95% CI, 2.3-3.5) in trials with run-in and 3.5 (95% CI, 1.2-10.7) in trials without run-in; for EPS, these ORs were 1.8 (95% CI, 1.4-2.2) and 2.0 (95% CI, 1.3-3.1) respectively, and for mortality 1.4 (95% CI, 1.0-2.0) and 1.6 (95% CI, 0.7-3.4). The use of placebo/investigated drug during run-in and run-in duration did not affect the estimates in a consistent way.

Conclusions: The use of run-in in trials might have led to overestimated efficacy and especially underestimated risks of side effects of antipsychotics compared with placebo for NPS in dementia.

Keywords
antipsychotics, dementia, efficacy, meta-analysis, pharmacoepidemiology, run-in, side effects
1 | INTRODUCTION

Results of randomized controlled trials are important for regulatory and clinical decisions. Researchers have therefore sought to optimize treatment effects and identify patients that will benefit most from treatment. One way of enhancing trial design is using a run-in period between screening for eligibility and before randomization. During this period of usually 1 to 2 weeks, drugs that the eligible patients already used are washed out. In some trials, the drugs are replaced by placebo to blind the participants for the change in treatment. Drug-naïve patients can also be given placebo, or the active drug of interest. Patients with high placebo response, poor compliance, low treatment response, or intolerance for the drug can thus be identified. At the end of the run-in phase, the researchers select the participants that are definitively included in the study. It is assumed that a run-in period will decrease placebo response and dropout during the trial and consequently increase the effect size and the power of a trial.

A small number of reviews have studied the effect that a run-in period can have on trial outcomes of psychopharmacological drugs. Antidepressants in children were 15% more effective in trials with a run-in period than in trials without a run-in period and above the threshold for a small effect size (standardized mean difference 0.26 vs 0.17, respectively; cut-off for small effect is 0.20). Another meta-analysis of antidepressant trials in depressed outpatients showed that a placebo run-in period was associated with higher efficacy and more power. On the other hand, run-in periods were not associated with greater efficacy in trials of antidepressants for major depression, benzodiazepines for anxiety, and naltrexon for alcohol addiction.

To our knowledge, the effect of a run-in period on efficacy and side effects of antipsychotics has not been investigated before. This is notable, because high placebo response rates, high dropout rates, and decreasing effectiveness over the years are the major problems in antipsychotic trials. An association between use of a run-in period and drug safety is not unlikely, because run-in periods can lead to exclusion of persons not tolerating the drugs and of noncompliant subjects. Moreover, atypical antipsychotics have been marketed with the claim of a more favorable side effects profile compared with conventional antipsychotics, i.e., lower rates of somnolence and extrapyramidal symptoms (EPS).

Antipsychotics are often prescribed for neuropsychiatric symptoms in dementia. Trials that tested the efficacy of antipsychotics for this indication commonly used run-in periods. The aim of this study was to assess the association of run-in periods in trials of conventional and atypical antipsychotics in dementia with clinical outcomes and also dropout.

2 | METHODS

We performed a meta-epidemiological study. We wrote a research proposal for the sponsor in advance, and it can be requested from the corresponding author.

Key Points

- Run-in periods have been used in the majority of trials that tested conventional and atypical antipsychotics in dementia.
- The use of run-in periods increased the estimated efficacy of antipsychotics in dementia.
- The use of run-in periods decreased the risk of somnolence, extrapyramidal symptoms, and mortality of antipsychotics in dementia.
- The use of run-in periods did not affect drop-out substantially.
- The effect of run-in on clinical outcomes of trials needs to be addressed as part of reviews.

2.1 | Search strategy

Four sources were used to identify trials. Two reviewers (T.A.H. and H.J.L.) first searched the electronic databases Cinahl, Embase, Pubmed, and Cochrane library with the strings "generic name atypical/conventional antipsychotic" AND trial AND dementia (see the Appendix S1). We composed a list of all conventional and atypical antipsychotics from the websites of the World Health Organization, Food and Drug Administration, and Wikipedia to enable this search. Secondly, we hand-searched the references of published systematic reviews, which were identified with the same electronic databases. Titles and abstracts of potentially eligible studies were retrieved in Pubmed. Thirdly, we sought RCTs in trial registration websites with the same keywords where possible. Finally, we searched the databases of the Dutch Medicines Evaluation Board and the FDA for unpublished trials of atypical antipsychotics.

2.2 | Study selection

Randomized placebo-controlled trials that tested the efficacy of orally administered conventional or atypical antipsychotics for neuropsychiatric symptoms in dementia were included. If studies seemed potentially eligible given title and abstract, full articles were retrieved as well as online protocols of unpublished studies. Two reviewers (T.A.H. and H.J.L.) reviewed these articles for definitive eligibility. Studies with no information on the use of a run-in period and with multiple drugs in one intervention arm were excluded. There were no restrictions with respect to publication date, language, flexible or fixed dosing of the active treatment, and duration of the study. The search was last rerun in June 2019.

2.3 | Data extraction

Two reviewers (T.A.H. and H.J.L.) independently extracted data from the included trials. First, we extracted general study characteristics: publication year, type of antipsychotic groups, setting, type of...
neuropsychiatric symptoms for which the antipsychotics were tested, and total number of randomized patients. Next, we extracted characteristics of the run-in period: presence/absence, replacement drug (no drug, placebo, or active treatment), and duration, as well as the percentage of patients excluded at the end of the run-in phase.

We then extracted the clinical outcomes for the drug and placebo groups. Efficacy of antipsychotics in dementia can be measured with a generic instrument that covers various neuropsychiatric symptoms (eg, NPI and BEHAVE-AD) or an instrument for specific symptoms such as agitation (eg, CMAI). We used the results measured with the instrument that matched the symptoms at enrollment, eg, if patients had to have agitation to enter a trial, we used the result reported with an agitation scale. We extracted the mean change from baseline to end point. When multiple dosages or multiple drug groups were included in a trial, an average change was calculated. We also extracted the standard deviation (SD) of the difference between the groups in mean change. If the SD was not reported, it was calculated with the P value, range, or confidence interval reported for the difference in mean change. Otherwise, the SD was imputed with the average of the reported SD of all trials with the same indication and instrument. For two trials that did not report the data we needed, we obtained the IPD and calculated the mean changes and SDs. In addition, we abstracted the number of patients with somnolence (sedation and drowsiness), and with EPS, and the number that died.

Finally, we extracted the total number of patients that dropped out (total drop-out) and the drop-out in the groups (selective drop-out). Run-in is often used to decrease drop-out and enhance power. Drop-out is also considered to represent the balance between efficacy and side effects.

The published main results’ article of a trial was our primary source of information. When the article did not report the data that we needed, secondary publications, trial reports, and meta-analyses-published online by industry were our secondary source. We contacted the authors of eligible trials to provide missing data, or individual patient data, and received such data of four studies.

The reviewers discussed differences in the extracted data until consensus was reached.

### 2.4 Data analyses

First, we assessed the relationship between the presence of a run-in period with the four clinical outcomes and selective dropout. We performed meta-analyses to pool efficacy, risk of somnolence, EPS, and mortality of the antipsychotic versus placebo groups in trials with and without run-in periods. For efficacy in terms of reduction in NPS, we calculated standardized mean differences (SMDs) to take into account the use of different instruments in the trials. SMDs were calculated with a 95% confidence interval (CI). For risk of somnolence, EPS, mortality, and selective dropout, we calculated odds ratios (ORs) with 95% CIs. Heterogeneity, presented as $I^2$, was calculated for all meta-analyses. A fixed-effects model was applied when $I^2$ was below 40%; otherwise a random-effects model.

We then investigated the relationship of the characteristics of the run-in periods with the clinical outcomes and selective dropout. The outcomes of trials that had run-in periods with and without a placebo were pooled (we did not find trials with active drugs in the run-in period). We also pooled outcomes of trials with a run-in period up to 1 week versus those with a longer run-in period (8 to 14 days). Finally, we tested the association between run-in characteristics and total drop-out with meta-regression.

We ran the above analyses for all antipsychotics combined first and then for the conventional antipsychotics and atypical antipsychotics separately. We wanted to perform an a priori sensitivity-analysis for atypical antipsychotics without quetiapine, because its efficacy and side effect profile are considered to differ. A post hoc analysis was run in which trials that had tested both an atypical drug and haloperidol were excluded from the analysis. All analyses were performed with STATA statistical software version 15.0.

### 3 RESULTS

Our search yielded 2768 potentially relevant RCTs (Figure 1). We obtained the reports of 65 RCTs for full text review. We identified 45 eligible RCTs, but four did not report whether a run-in period was used, four did not report any of the outcomes of interest, and two were ongoing. We used the other 35 studies in the current study.

Table 1 presents the general study characteristics. Twenty-nine of the 35 studies had a run-in period: nine of 11 conventional antipsychotic trials, 18 of 21 atypical antipsychotic trials, and two of three trials with both antipsychotics. Fourteen studies used placebo during the run-in period, two trials the investigated drug, and 19 studies no placebo or active treatment (washout only). The duration of the run-in periods varied between 2 days and 6 weeks. The percentage of patients excluded at the end of the run-in period varied from 0% to 29%. In five of seven relatively old conventional antipsychotic trials, the percentage was 0%, while in the atypical trials, it was at least 7%.

#### 3.1 Run-in and clinical outcome

The analysis of efficacy encompassed 29 of 35 studies (Table 2). The reduction in NPS in the drug versus placebo groups was somewhat higher in trials with a run-in period (SMD $-0.170$; 95% CI, $-0.227$ to $-0.112$) than in trials without a run-in period (SMD $-0.142$; 95% CI, $-0.331$ to 0.047). Efficacy was somewhat higher when placebo or active drug was used (SMD $-0.190$; 95% CI, $-0.267$ to $-0.112$), and when run-in lasted 1 week at most (SMD $-0.214$; 95% CI, $-0.289$ to $-0.138$).

The number of participants with somnolence during the study period was reported in 20 of 35 studies. The pooled risk for somnolence was lower when a run-in period was present (OR 2.8; 95% CI, 2.3-3.5) versus when it was absent (OR 3.5; 95% CI, 1.2-10.7). The use of placebo or active drug did not seem to affect the risk of
somnolence further. The risk was lower for trials with a run-in period of 1 week at most (OR 2.5; 95% CI, 1.7-3.9).

Nineteen of 35 studies reported the number of participants with EPS in the treatment groups. The risk of EPS in trials with a run-in period was slightly lower (OR 1.8; 95% CI, 1.4-2.2) than in trials without a run-in period (OR 2.0; 95% CI, 1.3-3.1). Use of placebo or duration of run-in did not seem to reduce the risk of EPS further.

The data of all but one trial could be used for the analysis of mortality risk. The risk of mortality was 1.6 (95% CI, 0.7-3.4) and 1.4 (95% CI, 1.0-2.0) in trials with and without run-in, respectively. The risk was slightly lower when no placebo was used, but run-in duration did not seem to affect it.

The sub-analyses in trials of atypical antipsychotics only yielded similar results for use of run-in versus no run-in on all outcomes (Table 2). In trials of conventional antipsychotics, however, efficacy seemed lower and risk of somnolence and EPS higher in trials with run-in compared to trials without run-in, but the number of trials without run-in was 3 at most and confidence intervals were large.

### 3.2 | Run-in and drop-out

Thirty-one of 35 studies reported dropout rates for the antipsychotic and placebo groups (Table 3). The number of participants that dropped out was 1519 of 4963 in the antipsychotic groups (30.6%) and 783 of 2747 in the placebo groups (28.5%). The odds ratio of selective dropout was 1.0 (95% CI, 0.9-1.2) for trials with a run-in period and 1.0 (95% CI, 0.7-1.3) for those without.

Total dropout in the studies varied between 0% and 81.7%. The use of a run-in period was not associated with a decreased total dropout (beta 0.3%; 95% CI, -15.1 to 15.7) (Table 3).

### 3.3 | Sensitivity analyses

There were not enough trials with and without run-in periods for the sensitivity-analysis of atypical antipsychotics without quetiapine. The results of the analysis without trials that tested both haloperidol and an atypical drug confirmed the pattern of higher efficacy and lower risk of side effects for the conventional and atypical antipsychotic drugs in trials with versus without run-in (Table S1).

### 4 | DISCUSSION

We assessed the association between use of a run-in period and clinical outcomes of 35 antipsychotic trials in dementia. The reduction in neuropsychiatric symptoms of antipsychotics versus placebo was somewhat higher in trials with a run-in period than in trials without a run-in period. The risk of somnolence, EPS, and mortality was lower in trials with than without run-in. Accordingly, the risk of dropout in the antipsychotic compared with the placebo groups, which represents the balance between beneficial and harmful effects, was not affected when run-in was used. Use of run-in periods did not influence total dropout either.

Several reviews have reported an association between the use of run-in periods and the efficacy of psychotropics. Two meta-analyses of antidepressant trials showed that a placebo run-in period was
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Antipsychotic</th>
<th>Setting</th>
<th>Neuropsychiatric Symptom</th>
<th>Run-in Period (Duration in Weeks)</th>
<th>Patients Excluded After Run-in, n (%)</th>
<th>Randomized Patients, n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hamilton, 1962</td>
<td>Trifluoperazine</td>
<td>Hospital</td>
<td>Diverse</td>
<td>No</td>
<td>NA</td>
<td>27</td>
</tr>
<tr>
<td>Sugerman, 1964</td>
<td>Haloperidol</td>
<td>Hospital</td>
<td>Diverse</td>
<td>No</td>
<td>NA</td>
<td>18</td>
</tr>
<tr>
<td>Rada, 1976</td>
<td>Thiopental</td>
<td>Hospital</td>
<td>Diverse</td>
<td>Yes, with placebo (1)</td>
<td>0 (0)</td>
<td>42</td>
</tr>
<tr>
<td>Barnes, 1982</td>
<td>Thoridazine</td>
<td>Nursing home</td>
<td>Diverse</td>
<td>Yes, with placebo (2)</td>
<td>7 (11.7)</td>
<td>53</td>
</tr>
<tr>
<td>Petrie, 1982</td>
<td>Haloperidol</td>
<td>Hospital</td>
<td>Diverse</td>
<td>Yes, with placebo (2)</td>
<td>0 (0)</td>
<td>61</td>
</tr>
<tr>
<td>Stotsky, 1984</td>
<td>Thoridazine</td>
<td>Nursing home and hospital</td>
<td>Diverse</td>
<td>Yes, washout only (2)</td>
<td>NR</td>
<td>358</td>
</tr>
<tr>
<td>Finkel, 1995</td>
<td>Haloperidol</td>
<td>Nursing home</td>
<td>Agitation</td>
<td>Yes, washout only (1)</td>
<td>0 (0)</td>
<td>33</td>
</tr>
<tr>
<td>Auchus, 1997</td>
<td>Haloperidol</td>
<td>Outpatients</td>
<td>Agitation</td>
<td>Yes, washout only (2)</td>
<td>0 (0)</td>
<td>12</td>
</tr>
<tr>
<td>Devanand, 1998</td>
<td>Haloperidol</td>
<td>Outpatients</td>
<td>Agitation</td>
<td>Yes, with placebo (1)</td>
<td>5 (7.0)</td>
<td>66</td>
</tr>
<tr>
<td>Teri, 2000</td>
<td>Haloperidol</td>
<td>Hospital</td>
<td>Agitation</td>
<td>Yes, washout only (2)</td>
<td>0 (0)</td>
<td>70</td>
</tr>
<tr>
<td>Pollock, 2002</td>
<td>Perphenazine</td>
<td>Nursing home</td>
<td>Diverse</td>
<td>Yes, washout only (+1)</td>
<td>NR</td>
<td>54</td>
</tr>
<tr>
<td>Satterlee, 1995</td>
<td>Olanzapine</td>
<td>Nursing home</td>
<td>Psychosis&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Yes, washout only</td>
<td>51 (17.7)</td>
<td>238</td>
</tr>
<tr>
<td>Janssen Ph, 1997</td>
<td>Risperidone</td>
<td>NR</td>
<td>Diverse</td>
<td>Yes, with placebo (1)</td>
<td>NR</td>
<td>39</td>
</tr>
<tr>
<td>Katz, 1999</td>
<td>Risperidone</td>
<td>Nursing home</td>
<td>Diverse</td>
<td>Yes, with placebo (1)</td>
<td>104 (14.3)</td>
<td>625</td>
</tr>
<tr>
<td>Street, 2000</td>
<td>Olanzapine</td>
<td>Nursing home</td>
<td>Diverse</td>
<td>Yes, with placebo (2)</td>
<td>82 (28.5)</td>
<td>206</td>
</tr>
<tr>
<td>Brodaty, 2003</td>
<td>Risperidone</td>
<td>Nursing home</td>
<td>Aggression</td>
<td>Yes, with placebo (1)</td>
<td>39 (10.2)</td>
<td>345</td>
</tr>
<tr>
<td>Janssen Ph, 2003</td>
<td>Risperidone</td>
<td>Nursing home</td>
<td>Psychosis</td>
<td>Yes, with placebo (1)</td>
<td>NR</td>
<td>18</td>
</tr>
<tr>
<td>De Deyn, 2004</td>
<td>Olanzapine</td>
<td>Nursing home</td>
<td>Psychosis</td>
<td>Yes, with placebo (2)</td>
<td>NR</td>
<td>652</td>
</tr>
<tr>
<td>Ballard, 2005</td>
<td>Quetiapine</td>
<td>Nursing home</td>
<td>Agitation</td>
<td>No</td>
<td>NA</td>
<td>62</td>
</tr>
<tr>
<td>De Deyn, 2005</td>
<td>Aripiprazole</td>
<td>Outpatients</td>
<td>Psychosis</td>
<td>Yes, washout only (1)</td>
<td>NR</td>
<td>208</td>
</tr>
<tr>
<td>Deberdt, 2005</td>
<td>Risperidone</td>
<td>Nursing home and outpatients</td>
<td>Psychosis</td>
<td>Yes, with placebo (2)</td>
<td>NR</td>
<td>494</td>
</tr>
<tr>
<td>Janssen Ph, 2005</td>
<td>Risperidone</td>
<td>NR</td>
<td>Psychosis</td>
<td>Yes, with risperidone (1)</td>
<td>NR</td>
<td>33</td>
</tr>
<tr>
<td>Mintzer, 2006</td>
<td>Risperidone</td>
<td>Nursing home</td>
<td>Psychosis</td>
<td>Yes, with placebo (1)</td>
<td>87 (15.5)</td>
<td>473</td>
</tr>
<tr>
<td>Schneider, 2006</td>
<td>Risperidone</td>
<td>Outpatients</td>
<td>Diverse</td>
<td>No</td>
<td>NA</td>
<td>421</td>
</tr>
<tr>
<td>Mintzer, 2007</td>
<td>Aripiprazole</td>
<td>Nursing home</td>
<td>Psychosis</td>
<td>Yes, washout only (1)</td>
<td>NR</td>
<td>487</td>
</tr>
<tr>
<td>Zhong, 2007</td>
<td>Quetiapine</td>
<td>Nursing home</td>
<td>Agitation</td>
<td>No</td>
<td>NA</td>
<td>333</td>
</tr>
<tr>
<td>Paleacu, 2008</td>
<td>Quetiapine</td>
<td>Not reported</td>
<td>Diverse</td>
<td>Yes, with quetiapine (2)</td>
<td>NR</td>
<td>40</td>
</tr>
<tr>
<td>Streim, 2008</td>
<td>Aripiprazole</td>
<td>Nursing home</td>
<td>Psychosis</td>
<td>Yes, washout only (1)</td>
<td>NR</td>
<td>256</td>
</tr>
<tr>
<td>Otsuka Ph, 2017a</td>
<td>Brexpiprazole</td>
<td>Nursing home</td>
<td>Agitation</td>
<td>Yes, washout only (6)</td>
<td>NR</td>
<td>413</td>
</tr>
<tr>
<td>Otsuka Ph, 2017b</td>
<td>Brexpiprazole</td>
<td>Nursing home and outpatients</td>
<td>Psychosis</td>
<td>Yes, washout only (6)</td>
<td>NR</td>
<td>270</td>
</tr>
<tr>
<td>ACADIA, 2018</td>
<td>Pimavanserin</td>
<td>Nursing home and outpatients</td>
<td>Agitation</td>
<td>Yes, washout only (4)</td>
<td>NR</td>
<td>111</td>
</tr>
<tr>
<td>Ballard, 2018</td>
<td>Pimavanserin</td>
<td>Nursing home</td>
<td>Psychosis</td>
<td>Yes, washout only (3)</td>
<td>25 (12.1)</td>
<td>181</td>
</tr>
<tr>
<td>Allain, 2000</td>
<td>Tiapride</td>
<td>Nursing home and hospital</td>
<td>Agitation</td>
<td>No</td>
<td>NA</td>
<td>306</td>
</tr>
<tr>
<td>De Deyn, 1999</td>
<td>Risperidone</td>
<td>Nursing home</td>
<td>Diverse</td>
<td>Yes, with placebo (1)</td>
<td>27 (7.3)</td>
<td>344</td>
</tr>
<tr>
<td>Tariot, 2006</td>
<td>Quetiapine</td>
<td>Nursing home</td>
<td>Psychosis</td>
<td>Yes, washout only (1)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>123 (24.6)</td>
<td>284</td>
</tr>
</tbody>
</table>

Abbreviations: NA not applicable; NR, not reported; Ph, pharmaceuticals.

<sup>a</sup>Reduction measured with generic instrument (in all other studies indication and outcome scale were congruent).

<sup>b</sup>At least 2 days.
Abbreviations: EPS, extrapyramidal symptoms; OR, odds ratio; SMD, standardized mean difference.

In observational studies, bias due to (de) selection of patients based on prior treatment and its outcomes, whether before or after the start of the study, is generally called selection bias.59

A common argument for use of run-in periods is the reduction of noncompliance and dropout. Our findings showed that a placebo run-in was not associated with lower between-group or total dropout rate.

### 4.1 Bias due to run-in periods

In the studies that we identified, patients that met the inclusion criteria could have been excluded from trial participation as a result of the outcomes during the run-in period. It is therefore not surprising that the use of run-in period yielded higher efficacy estimates and lower risks of side effects.3,58 In observational studies, bias due to (de) selection of patients based on prior treatment and its outcomes, whether before or after the start of the study, is generally called selection bias.59

Bias due to run-in periods in trials is not commonly discussed in the literature. Most tools for risk of bias assessment in trials do not require consideration of run-in periods either. In eight of 11 meta-analyses of antipsychotic trials in dementia, more than 80% of the included trials used a run-in period (Table S2),60-70 Risperidone and olanzapine, currently the two most popular antipsychotic drugs for use in dementia, have been tested in ten trials of which nine included a run-in period. The general assumption is that selection of patients before randomization is said to reduce only generalizability of study results, not the internal validity of the trial results.1

We propose a different view. The screening and selection of patients before randomization should be based on (contra-)indications that are applied in daily medical practice. The selected patients will then represent the patients of interest to doctors and the population of interest as defined in the PICOs of reviews. This selection needs to be distinguished from deselection of eligible patients based on observed treatment effects during run-in (between screening and randomization).2 The remaining randomized group does not represent the population of interest any more. Estimates of efficacy and risk of side effects will be biased for the population of interest. Therefore, run-in needs to be considered as a source of bias in trials and reviews of trials.

Another issue to consider is the ethics of entering patients who are doing (relatively) well on a certain antipsychotic drug into a trial of another or the same antipsychotic. During washout, symptoms could return, and it is questionable whether the patient will respond as favorably to a new drug.

### 4.2 Strengths and limitations

To our knowledge, this is the first study that investigated the relationship of run-in periods with clinical outcomes of antipsychotic trials including side effects. Our pooled estimates of efficacy seemed low but are corroborated by previous reviews reporting SMDs between
The authors received a grant from the Dutch Medicines Evaluation Board (College ter Beoordeling van Geneesmiddelen).

**ACKNOWLEDGEMENT**

The authors received a grant from the Dutch Medicines Evaluation Board (College ter Beoordeling van Geneesmiddelen).

### REFERENCES


### TABLE 3  Use of run-in periods and dropout in antipsychotic trials in dementia

<table>
<thead>
<tr>
<th>Drug group</th>
<th>Total Dropout</th>
<th>Selective Dropout</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (95% CI) N</td>
<td>Beta (95% CI) N</td>
</tr>
<tr>
<td>Conventional and atypical antipsychotics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No run-in</td>
<td>1.0 (0.7-1.3) 5 ref</td>
<td></td>
</tr>
<tr>
<td>With run-in</td>
<td>1.0 (0.9-1.2) 26</td>
<td>0.3 (−15.1 to 15.7) 33</td>
</tr>
<tr>
<td>washout only</td>
<td>0.9 (0.7-1.1) 13 ref</td>
<td></td>
</tr>
<tr>
<td>With placebo/ drug</td>
<td>1.2 (1.0-1.4) 13</td>
<td>−4.7 (−14.9 to 5.5) 27</td>
</tr>
<tr>
<td>duration ≤ 1 week</td>
<td>0.9 (0.8-1.1) 14 ref</td>
<td></td>
</tr>
<tr>
<td>duration &gt; 1 week</td>
<td>1.2 (1.0-1.5) 12</td>
<td>−5.0 (−14.7 to 4.6) 26</td>
</tr>
<tr>
<td>Atypical antipsychotics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No run-in</td>
<td>1.4 (0.7-2.8) 2 ref</td>
<td></td>
</tr>
<tr>
<td>With run-in</td>
<td>1.0 (0.7-1.3) 9</td>
<td>16.0 (−7.5 to 39.6) 13</td>
</tr>
<tr>
<td>With run-in</td>
<td>1.0 (0.9-1.2) 19</td>
<td>−7.7 (−25.1 to 9.7) 23</td>
</tr>
</tbody>
</table>

AAbbreviation: OR, odds ratio.

*Drug group versus placebo.

*A random effects model was used.

0.12 and 0.21,61,66,70 SMDs above the threshold of 0.200 for a small treatment effect were mainly found in meta-analyses that focused on aggression or agitation.63,64,68,70

A limitation of our research was that only six studies did not use a run-in phase. Most of these studies were performed in outpatients and with atypical antipsychotics, in particular quetiapine. As a result of this distribution, the higher efficacy in trials with run-in might be partly attributable to a higher efficacy of conventional antipsychotics. Nevertheless, our sensitivity analysis in atypical antipsychotic trials showed a higher efficacy for trials with a run-in period as well. Additionally, one would expect the risk of side effects to increase with run-in as well, but it did not.

### 5 CONCLUSION

The use of a run-in period is very common in antipsychotic trials for dementia. In these trials, efficacy was higher compared with trials without run-in, while the risk of side effects was lower. Therefore, the use of a run-in period in trials might have led to overestimated efficacy and especially underestimated side effects of antipsychotics for neuropsychiatric symptoms in dementia. Meta-analyses should include sensitivity-analyses of trials with and without run-in periods.

### AUTHOR CONTRIBUTIONS

T.A. Hulshof and H.J. Luijendijk searched and selected the trials, extracted the data, and drafted the manuscript. T.A. Hulshof performed the data-analysis. H.J. Luijendijk designed the study. S.U. Zuidema and C.C. Gispen-de Wied critically commented on the design and results of the study. All authors reviewed the manuscript and suggested revisions.

### ORCID

Hendrika J. Luijendijk https://orcid.org/0000-0002-6865-0128

SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section at the end of this article.