Received: 24 April 2018 | Revised: 11 September 2018 | Accepted: 17 September 2018

DOI: 10.1002/mpr.1757

ORIGINAL ARTICLE

WILEY

Baseline imbalances and clinical outcomes of atypical antipsychotics in dementia: A meta-epidemiological study of randomized trials

Tessa A. Hulshof[1] | Sytse U. Zuidema[1] | Peter J.K. van Meer[2] | Christine C. Gispen-de Wied[2] | Hendrika J. Luijendijk[1]

[1] University Medical Center Groningen, Department of General Practice and Elderly Care Medicine, University of Groningen, Groningen, The Netherlands

[2] Medicines Evaluation Board, Utrecht, The Netherlands

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

Funding information
Dutch Medicines Evaluation Board

Abstract

Objectives: To assess baseline imbalances in placebo-controlled trials of atypical antipsychotics in dementia, and their association with neuropsychiatric symptoms (NPS), extrapyramidal symptoms (EPS), and mortality.

Method: We searched for trials in multiple sources. Two reviewers extracted baseline characteristics and outcomes per treatment group. We calculated direction, range, pooled mean, and heterogeneity in the baseline differences, and used meta-regression for the relationship with the outcomes.

Results: We identified 23 trials. Baseline type of dementia, cognitive impairment and NPS were poorly reported. The drug group had a higher mean age than the placebo group in nine trials and lower mean age in three trials (p = 0.073). The difference in percentage men between the drug and placebo group ranged from −9.7% to 4.4%. There were no statistically significant pooled baseline differences, but heterogeneity was present for age. Higher mean age at baseline in the drug versus placebo group was significantly associated with greater reduction in NPS, and higher percentage of non-White persons with lower risk of EPS. Imbalances were not significantly associated with risk of mortality.

Conclusion: Randomized trials of atypical antipsychotics in dementia showed baseline imbalances that were associated with higher efficacy and lower risk of EPS for atypical antipsychotics versus placebo.

KEYWORDS

antipsychotics, baseline imbalances, bias, meta-regression, trials

1 | INTRODUCTION

Randomization is the cornerstone of clinical trials. Randomization is used to ensure that chance instead of patient characteristics determine treatment assignment. In daily medical practice, patient characteristics affect the choice of treatment, and will therefore be distributed unevenly across treatment groups. Moreover, these patient characteristics may also affect prognosis independently of the effect of treatment. Hence, treatment groups of a trial need to be comparable to establish an unbiased effect of a treatment on prognosis (clinical outcomes).

The larger the sample size of a randomized trial is, the smaller the differences between treatment groups at the start of the trial will become and the more comparable the groups will be. Yet, imbalanced groups can occur by chance despite adequately designed and conducted randomization procedures, especially in trials with small sample sizes. In addition, flawed or corrupted randomization procedures can give rise to systematic baseline imbalances between groups (Clark et al., 2014; Hróbjartsson, Boutron, Turner, Altman, & Moher, 2013). If the baseline imbalances are distributed in the same way across trials testing the same treatment, they will bias the pooled results of those trials too (Corbett, Higgins, & Woolacott, 2014; Egger, Juni, Bartlett, Holenstein, & Sterne, 2003; Luijendijk & Hulshof, 2015; Riley et al., 2013; Schulz, Chalmers, Hayes, & Altman, 1995; Trowman, Dumville, Torgerson, & Cranny, 2007).

In a previous meta-analysis, we observed baseline imbalances in trials testing antipsychotic drugs in dementia (Hulshof, Zuidema, Ostelo, & Luijendijk, 2015). Atypical antipsychotic have been found to reduce neuropsychiatric symptoms (NPS) and to increase the risk of extrapyramidal symptoms (EPS) and mortality (Schneider, Dagerman, & Insel, 2005; Ma et al., 2014). In some trials of atypical antipsychotics, the baseline characteristics of the atypical group seemed unfavorable in comparison with placebo (Ballard et al., 2005; Street, Clark, Gannon, Cummings, & Bymaster, 2000; Zhong, Tariot, Mintzer, Minkwitz, & Devine, 2007): Patients were older, and more often men or diagnosed with vascular dementia. These factors are predictive of EPS and death in patients with dementia (Fitzpatrick, Kuller, Lopez, Kawas, & Jagust, 2005; Gambassi et al., 1999; Garcia-Ptacek et al., 2014; Mitchell et al., 2004; Rountree, Chan, Pavlik, Darby, & Doody, 2012). Therefore, not the atypical drugs themselves but the vulnerability of the patients in the drug compared with the placebo groups could have resulted in a higher risk of EPS and deaths. Moreover, if the more vulnerable patients had more severe NPS and dropped out more often, the remaining group of patients would have had less NPS. Consequently, the effect of atypical antipsychotics on symptom reduction might have overestimated (Hernán, Hernández-Díaz, & Robins, 2004).

Atypical antipsychotics were introduced to the market with the claim that these drug were as effective as haloperidol but had less side effects (Allain et al., 2000; De Deyn, Rabheru, & Rasmussen, 1999; Tariot et al., 2006). At the time, haloperidol—a typical antipsychotic—was the first choice of treatment for agitation and psychosis in dementia. To substantiate the claim of relative benefits and harms, atypical antipsychotics and haloperidol have been compared with placebo simultaneously in a number of trials. We observed that in trials with an haloperidol group, the atypical antipsychotics groups seemed to be less vulnerable than the placebo groups (Allain et al., 2000; De Deyn et al., 1999; Hulshof et al., 2015; Tariot et al., 2006). This imbalance might have led to overestimation of the effect of atypical antipsychotics on reduction of NPS and underestimation of the risk of EPS compared with placebo. The variation in baseline imbalances between atypical drug groups and placebo groups across trials enables an evaluation of the effect that the imbalances might have had on trial results.

The aim of this study was (a) to assess the presence of systematic baseline imbalances in placebo-controlled trials of atypical antipsychotics in dementia, and (b) to evaluate the association of baseline imbalances with reduction of NPS and risk of EPS and mortality.

2 | METHODS

2.1 | Literature search and selection

Two reviewers (T. A. H. and H. J. L.) searched trials in four sources. First, we used the electronic databases Pubmed, Cinahl, and Embase and entered the strings ("generic name atypical antipsychotic" AND trial) and (dementia). We had composed a list of all atypical antipsychotics from the websites of the World Health Organization and the Food and Drug Administration to this end (US Food and Drug Administration, 2013; World Health Organization, 2013). Second, we hand-searched the references of published systematic reviews, which were identified with the same electronic databases and the Cochrane library. Titles and abstracts of potentially eligible studies were retrieved in the Pubmed. Third, we sought randomised controlled trials (RCTs) in trial registration websites with the keywords ("new generation", "second generation," or "atypical") and "antipsychotic". Finally, we searched the databases of the Dutch Medicines Evaluation Board and the FDA for unpublished trials.

If studies seemed potentially eligible given title and abstract, full articles of published studies 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. Randomized placebo-controlled trials that reported the effect of orally administered atypical antipsychotics on NPS or mortality in participants with Alzheimer's disease or vascular dementia were included. Studies in patients with Lewy body or Parkinson's dementia were excluded, because they are much less tolerant for antipsychotics, as were studies with multiple drugs in one intervention arm. 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 August 2017.

2.2 | Data extraction

Two reviewers (T. A. H. and H. J. L.) independently extracted data from the included studies. First, we extracted general study characteristics: setting, type of dementia, comparison groups, study duration, number of randomized patients in each treatment group, and publication status (published full article, or unpublished) and commercial funding. We assessed the randomization procedures consisting of the random sequence generation and allocation concealment, defined and scored as having a low, unclear or high risk of bias in accordance with the Cochrane Risk of Bias assessment tool (Higgins, Altman, & Sterne, 2014). We also recorded whether information about baseline characteristics of the treatment group for *all* randomized patients was presented in a baseline table in line with CONSORT requirement (Schulz, Altman, & Moher, 2010).

Secondly, we extracted the baseline characteristics of *all* randomized participants in the atypical antipsychotic and placebo groups: mean age and standard deviation (*SD*), number of men, number of non-White persons, number of vascular/mixed dementia, mean severity of dementia and *SD*, mean severity of NPS and *SD*. For severity of dementia, we used the mean Mini-Mental State Examination (MMSE) score because it was the most frequently reported instrument (eTable3). For severity of NPS, we recorded the reported mean NPI(-NH), BEHAVE-AD, BPRS, BRSD, or neurobehavioral rating scale score. In case of multiple

reported generic instruments, we preferred the most commonly used NPI(-NH), or otherwise the BEHAVE-AD. Other potentially important prognostic baseline characteristics, such as (cardiovascular) comorbidity, somatic and psychiatric medication use, and EPS were reported too infrequently to be of use for our analyses (eTable1). When studies included multiple active medication groups (different dosages or drugs), an average mean or percentage was calculated for these groups. If the *SD* for mean age, MMSE, and NPS was missing, the *SD* was imputed with the average *SD* of the other trials.

Finally, we extracted the clinical outcomes. Efficacy of antipsychotics in dementia is usually measured with a generic instrument for diverse NPS (e.g., NPI and BEHAVE-AD) or with an instrument specific for one type of symptoms such as aggression (CMAI). We preferred the reported total score of a generic instrument to guarantee the comparability of outcomes across trials, but if it was lacking, we will used the reported total score of the specific instrument. If multiple generic instruments were used, we extracted the most commonly reported (NPI(-NH) or otherwise BEHAVE-AD). We extracted the mean change from baseline to end point with its *SD* for the active drug and placebo groups. If the confidence interval, standard error, or *p* value was reported, the *SD* was calculated with this information (Cochrane handbook). When multiple dosage or multiple drug groups were included in a trial, an average change was calculated. We also recorded the number of patients with EPS and the number of deaths during the trial.

For all extracted information, the published article of a trial was our primary source. Authors provided additional information at our request (De Deyn et al., 1999; De Deyn, Jeste, Swanink, Kostic, & Breder, 2005; Kurlan, Cummings, Raman, & Thal, 2007; Paleacu, Barak, Mirecky, & Mazeh, 2008; Schneider et al., 2006) and meta-analyses published by industry were our secondary source. Other articles and meta-analyses were our tertiary source of information (Ballard, Waite, & Birks, 2006; Carson, McDonagh, & Peterson, 2006; Cheung & Stapelberg, 2011; Lee et al., 2004; Lonergan, Luxenberg, Colford, & Birks, 2010; Maher et al., 2011; Schneider, Dagerman, & Insel, 2006; Seitz et al., 2013). The reviewers discussed the differences in the extracted data until consensus was reached.

2.3 | Statistical analyses

First, we plotted the difference between group sizes (drug vs. placebo) against total trial size for 17 trials with unrestricted randomization (eTable1), and the expected distributions for the 50% and 95% prediction intervals (Schulz, Chalmers, Grimes, & Altman, 1994). For trials with more than two active drug groups, we used the first reported active drug group (and a total trial size of placebo group plus first active drug group). For studies that used a randomization ratio other than 1:1 for placebo versus active drug group, we recalculated the size of the active drug group by dividing the true size by the inverse of the ratio, and then recalculated the hence found difference back to the original total trial size. We then plotted this difference against to the true total number of participants. Trials that reported blocked randomization were excluded from this analysis.

Second, we described the range and direction of the baseline imbalances for studies with and without haloperidol groups. The rationale for this distinction is that studies with a haloperidol group seem to suggest higher efficacy, lower risk of EPS, and lower risk of mortality than trials without a haloperidol group (see eFigure 1–3). We then computed a one-sided sign-test per characteristic to test whether the proportion of studies that reported an imbalance in the most common direction (e.g., higher mean age in antipsychotic vs. placebo group) was higher than can be expected by chance (50%). Studies that reported no difference between groups (e.g., same mean age, which could be due to rounding) and studies with a missing baseline difference are automatically discarded from a sign-test.

Third, we performed meta-analyses to calculate the pooled mean difference (MD) for baseline age, severity of dementia and severity of NPS, and the pooled odds ratio (OR) for men, non-White persons, and vascular/mixed dementia with fixed-effects models (Clark, Fairhurst, Cook, & Torgerson, 2015; Ebell, 2013; Trowman et al., 2007). We expected a common effect estimate of zero in these mean baseline variables. Again, we distinguished between studies with and without haloperidol groups. The analyses generate an I^2-statistic for heterogeneity. We calculated 95% confidence intervals around I^2 with the direct command `heterogi` in Stata.

Fourth, we performed a meta-regression analyses to assess the relationship of the individual baseline imbalances for *all* randomized patients with reduction in NPS, risk of EPS, and risk of mortality. The beta-coefficients (betas) were calculated with 95% confidence intervals. We estimated the standardized mean differences (SMD) for NPS reduction and OR for risk of EPS and mortality. As many trials reported no deaths in one or both treatment groups, we used the Mantel–Haenszel weighted fixed effects model with continuity correction based on the reciprocal of the opposite group arm size to calculate the pooled ORs (Sweeting, Sutton, & Lambert, 2004). A fixed-effects model was applied when heterogeneity (I^2) was found to be below 40%, otherwise a random-effects model (DerSimonian and Laird model with the estimate of heterogeneity being taken from the from the Mantel–Haenszel model) (Deeks, Higgins, & Altman, 2011). The plot of group size difference against total trial size was made in R (R CoreTeam, 2013), the other analyses were performed with Stata version 14.1 (StataCorp., 2013).

When we found that a large number of baseline differences were not reported, we decided to pool the outcomes of studies reporting and not reporting a baseline characteristic in a post hoc analysis. Given the discrepancy in results of trials with and without haloperidol group, this analysis was restricted to the latter type of trials.

3 | RESULTS

Our search yielded 1,997 potentially relevant RCTs (Figure 1). We obtained the reports of 29 RCTs for full text review and finally identified 23 eligible RCTs with 5,853 participants (Allain et al., 2000; Brodaty et al., 2003; De Deyn et al., 2004, 1999; Deberdt et al., 2005; Herz, Volicer, Frankenburg, Colon, & Kittur, 2002; ILO522, 2002; Katz et al., 1999; Kennedy et al., 2005; Mintzer et al., 2006; Mintzer, Tune, Breder, Swanink, & Marcus, 2007; Paleacu et al., 2008; RIS-BEL-14, 1997; RIS-INT-83, 1997; Satterlee, 1995; Schneider, Tariot, et al., 2006; Street et al., 2000; Streim, Porsteinsson, Breder, Swanink, & Marcus, 2008; Tariot et al., 2006; Zhong et al., 2007; ZIP-128-105, 1993). Five trials were relatively small and

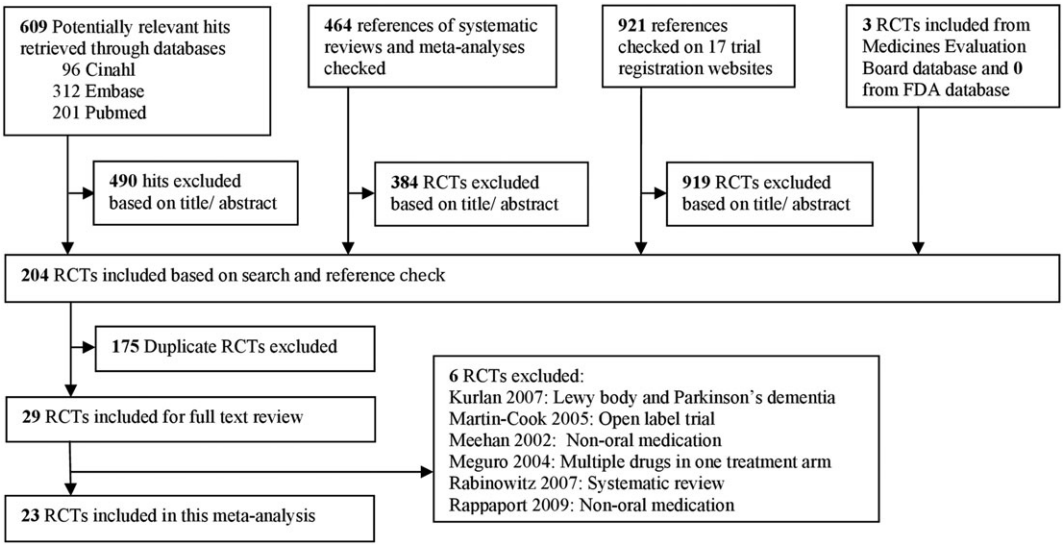

FIGURE 1 Flow diagram of literature search and study selection

TABLE 1 Characteristics of randomized placebo-controlled trials that tested atypical antipsychotics in dementia

Author and Year	Antipsychotic drug	Setting	Type of dementia	*N* randomized	Duration, weeks	Dose and range in mg/d	Published[b]
ZIP-128-105, 1993	Ziprasidone	Nursing home	AD-VAS	23	4	2–6	No
Satterlee, 1995	Olanzapine	Nursing home	AD	238	8	1–8	No
Ris-Bel-14, 1997[a]	Risperidone	NR	AD	39	4	1–4	No
Ris-Int-83, 1997[a]	Risperidone	NR	AD	18	8	0.5–1.5	No
De Deyn, 1999	Risperidone Haloperidol	Nursing home	AD-VAS-MIX	344	12	0.5–4 0.5–4	Yes
Katz, 1999	Risperidone	Nursing home	AD-VAS-MIX	625	12	0.5, 1, 2[c]	Yes
Allain, 2000	Tiapride Haloperidol	Nursing home & Hospital	AD-VAS-MIX	306	3	100–300 2–6	Yes
Street, 2000	Olanzapine	Nursing home	AD	206	6	5, 10, 15[c]	Yes
Herz, 2002	Risperidone Olanzapine	NR	AD	29	10	0.5–4 2.5–20	No
ILO522, 2002	Iloperidone	NR	AD-VAS-MIX	15	4	0.5–6	No
Brodaty, 2003	Risperidone	Nursing home	AD-VAS-MIX	345	12	0.25–2	Yes
De Deyn, 2004	Olanzapine	Nursing home	AD	652	10	1, 2.5, 5, 7.5[c]	Yes
Ballard, 2005	Quetiapine	Nursing home	AD	62	6	50–100	Yes
De Deyn, 2005	Aripiprazole	Outpatients	AD	208	10	2–15	Yes
Deberdt, 2005	Risperidone Olanzapine	Nursing home & Outpatients	AD-VAS-MIX	494	10	0.5–2 2.5–10	Yes
Kennedy, 2005	Olanzapine	Outpatients	AD (no NPS)	268	26	2.5–7.5	Yes
Mintzer, 2006	Risperidone	Nursing home	AD-VAS	473	8	0.5–1.5	Yes
Schneider, 2006	Risperidone Olanzapine Quetiapine	Outpatients	AD	421	2–36	0–2[d] 0–17.5 0–200	Yes
Tariot, 2006	Quetiapine Haloperidol	Nursing home	AD	284	10	25–600 0.5–12	Yes
Mintzer, 2007	Aripiprazole	Nursing home	AD	487	10	2, 5, 10[c]	Yes
Zhong, 2007	Quetiapine	Nursing home	AD-VAS	333	10	100, 200[c]	Yes
Paleacu, 2008	Quetiapine	NR	AD	40	6	75–300	Yes
Streim, 2008	Aripiprazole	Nursing home	AD	256	10	0.7–15	Yes

Note. AD: Alzheimer disease; NPS: neuropsychiatric symptoms; NR: not reported; VAS: vascular dementia; Mix; mixed dementia.

[a]mortality data were published in Haupt, 2006(Haupt, Cruz-Jentoft, & Jeste, 2006).

[b]Trial with conference abstracts only were considered as unpublished.

[c]Groups.

[d]Doctors were allowed to stop medication if deemed inefficient or causing too much side-effects.

unpublished (Herz et al., 2002; ILO522, 2002; RIS-BEL-14, 1997; RIS-INT-83, 1997; ZIP-128-105, 1993; Table 1). Twenty trials investigated one atypical antipsychotic drug, three of which included an extra haloperidol group, (Allain et al., 2000; De Deyn et al., 1999; Tariot et al., 2006) and three trials investigated multiple atypical drugs. The follow-up was ≤12 weeks in 22 trials, and ≥26 weeks in two trials. All trials were sponsored completely or partly by industry; one trial did not report the source of funding (Herz et al., 2002).

No study described the randomization procedure completely in terms of both the random sequence generation and allocation concealment (eTable 1 and eTable 2). Baseline characteristics were also poorly reported. Only 13 studies presented a baseline table or baseline information in the text for all randomized patients, two studies for a selection of all randomized patients, and eight studies, including four published studies, did not present a baseline table or baseline information in the text (Etable3). Only three trials reported all six patients characteristics. The first author of two trials provided additional data (De Deyn et al., 2005, 1999). For another trial, we calculated the missing baseline information with the provided IPD (Paleacu et al., 2008).

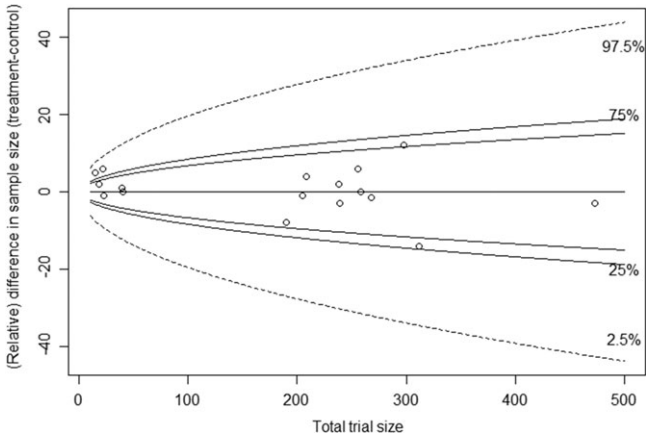

FIGURE 2 Treatment versus control group size differences and total trial size for 17 unrestricted trials, with expected distributions for 50% and 95% (dotted) prediction intervals

Figure 2 presents the relation between the difference between group sizes (placebo vs. active) and the total sample size for 17 trials with unrestricted randomization, together with the expected distributions for the 50% and 95% prediction intervals. Less than expected trials were outside the 50% and 95% distribution lines, four (24%) and zero (0%) respectively.

Table 2 shows the range of the actual differences between the placebo and treatment groups for each of the baseline characteristics. The percentage men and vascular/mixed dementia showed imbalances ranging from −9.7% to 4.4% and −9.9% to 2.7% respectively. One trial showed a difference in NPS of 0.271. Baseline imbalances that we investigated were not accounted for in the analyses of all but two trials (Kurlan et al., 2007; Schneider, Tariot, et al., 2006).

Table 2 also shows the direction of the baseline imbalances. No statistical differences were found. When there was no haloperidol group, there were numerically more trials with a higher age in the antipsychotics versus placebo group (8 vs. 2; p = 0.055), with a higher percentage of men (7 vs. 5; p = 0.387) and a higher percentage of vascular/mixed dementia (8 vs. 3; p = 0.113). When combining all trials, the number of trials with a positive versus negative direction was numerically higher for age (9 vs. 3; p = 0.073) and men (10 vs. 5; p = 0.125), and lower for severity of NPS (1 vs. 4; p = 0.188).

Table 3 presents the size and heterogeneity of the pooled baseline differences. The pooled imbalance in the percentage of men in the trials with a haloperidol group stood out numerically (5.4%), but none of the imbalances were statistically significantly different. Four of six baseline characteristics exhibited heterogeneity, when there

TABLE 2 Range and direction of baseline differences between atypical antipsychotic and placebo groups

Patient characteristic	Trials with reported characteristic or IPD for all randomized, k/n	Trials without haloperidol group			Trials with haloperidol group			All trials
		Difference, range[a]	Trials, $n-/n0/n+$ [d]	Sign test, p[e]	Difference, range[a]	Trials, $n-/n0/n+$ [d]	Sign test, p[e]	Sign test, p[e]
Age in years, mean	17/23	−1.9 to 2.0	2/4/8	0.055	−2.0 to 1.7	1/1/1	0.750	0.073
Male gender, %	15/21[f]	−9.7 to 4.4	5/0/7	0.387	2.3 to 7.3	0/0/3	0.125	0.151
Non-White race, %	13/21[f]	−4.9 to 3.1	3/0/8	0.113	−0.9 to −0.0	2/0/0	0.250	0.291
Vascular or mixed dementia, %	5/13	−3.3 to 2.7	1/0/3	0.313	−9.9 (1 trial)	1/0/0	0.500	0.500
MMSE, mean[b]	8/23	−1.2 to 0.7	2/1/4	0.344	−0.2 (1 trial)	1/0/0	0.500	0.500
NPS, standardized mean[c]	5/22[f]	−0.146 to 0.271	2/0/1	0.500	−0.126 to −0.048	2/0/0	0.250	0s.188

Note. IPD: individual patient data; MMSE: mini-mental state examination; NPS: neuropsychiatric symptoms.

[a]the baseline difference for each trial was calculated as the mean or percentage in the atypical antipsychotic group minus the mean or percentage of the placebo group.

[b]MMSE can be scored between 0 and 30; higher is better.

[c]NPS were measured with different instruments in the studies and therefore the mean was standardized with the *SD*; higher is worse.

[d]$n-$ stands for the number of trials with a negative baseline imbalance (f.i. lower age in antipsychotic versus placebo group), and $n0$ for the number of trials with no baseline imbalance (f.i. similar mean age in antipsychotic and placebo group), $n+$ for the number of trials with a positive baseline imbalance (f.i. higher age in antipsychotic versus placebo group).

[e]one-sided sign-test per characteristic to test whether the proportion of studies that reported an imbalance in the direction (in the most common direction) could be attributed to chance.

[f]less than 23, because two trials were performed in men only, two trials in White persons only, 10 trials in patients with Alzheimer disease only, and one trial in patients without NPS.

TABLE 3 Pooled baseline difference and heterogeneity in atypical antipsychotic versus placebo groups

	Trials without haloperidol group		Trials with haloperidol group		All trials	
Patient characteristic	Pooled difference (95% CI)	I^2, % (95% CI)	Pooled difference (95%CI)	I^2, % (95% CI)	Pooled difference (95% CI)	I^2, % (95% CI)
Age in years, mean	0.1 (−0.4; 0.6)	0 (0–55)	−0.2 (−1.3; 1.0)	70 (0–91)	0.1 (−0.4; 0.5)	12 (0–49)
Male gender, %	0.3 (−2.8; 3.4)	0 (0–58)	5.4 (−1.9; 12.7)	0 (0–90)	1.1 (−1.8; 3.9)	0 (0–54)
Non-White race, %	−0.1 (−2.5; 2.3)	0 (0–60)	−0.5 (−2.3; 2.0)	0 (nt)	−0.1 (−2.3; 2.0)	0 (0–57)
Vascular or mixed dementia, %	0.3 (−3.7; 4.2)	0 (0–85)	−9.9 (−21.8; 2.0)	nt (nt)	−1.0 (−4.8; 2.7)	12 (0–82)
MMSE, mean[a]	0.1 (−0.4; 0.5)	9 (0–71)	−0.2 (−1.6; 1.2)	nt (nt)	0.0 (−0.4; 0.5)	5 (0–68)
NPS, standardized mean[b]	−0.120 (−.252; .011)	27 (0–92)	0.013 (−.178; .205)	0 (nt)	−0.077 (−.186; .031)	11 (0–81)

Note. CI: confidence interval; MMSE: mini-mental state examination; NPS: neuropsychiatric symptoms; nt: not testable (too few studies).

[a]MMSE can be scored between 0 and 30; higher is better.

[b]NPS were measured with different instruments in the studies and therefore the mean was standardized with the *SD*; higher worse.

should have been none. Heterogeneity was 70% for the baseline difference in age in the trials with haloperidol group, and 27% for the difference in severity of NPS in trials without haloperidol group. None of the confidence intervals around I^2 suggested statistically significant heterogeneity.

Table 4 presents the associations between individual baseline imbalances and the clinical outcomes. Only for age, sex, and race were there more than 10 trials, the minimum for a reliable meta-regression analysis. A higher mean age, a higher percentage of men, and of persons of non-White race in the atypical antipsychotic drug than the placebo group, which was more often the case than not (see Table 2), was associated with greater efficacy and lower risk of EPS. In particular, one percentage more males in the treatment versus placebo group was statistically significantly associated with a higher reduction in NPS (beta −0.027; 95% CI [−0.047, −0.006]), and one percentage more non-White persons with a lower risk of EPS (beta −0.4; 95% CI [−0.8, −0.1]). An association with mortality risk could be not confirmed for any of the baseline imbalances.

As half of the baseline imbalances we wanted to abstract were not reported, we pooled the clinical outcomes for trials with and without missing baseline information for each baseline characteristic. Efficacy was consistently higher and risk of EPS consistently lower in studies without baseline information than for studies with this information (Table 5). Risk of mortality was, however, lower in studies with missing age, sex, and type of dementia, but higher in studies with missing race, MMSE, and severity of NPS.

4 | DISCUSSION

We reviewed the randomization procedures and baseline imbalances of 23 randomized placebo-controlled trials of atypical antipsychotics in 5,853 patients with dementia. All trials reported the randomization procedures incompletely, and only three trials reported the six baseline characteristics of interest for all randomized patients. Numerically more trials reported a higher mean age and a higher percentage of men and of non-White persons in the atypical antipsychotics group than in the placebo group. These imbalances were associated with greater efficacy and lower risk of EPS, but not with risk of mortality. Trials with missing baseline information seemed to have a more favorable pooled efficacy and lower risk of EPS than trials that reported this information.

4.1 | Randomization procedures

The goal of random sequence generation and concealment of allocation is that investigators, physicians, and patients cannot foresee allocation and then change the decision or time to enroll, or change the allocation itself. If executed correctly, randomization will distribute measured and unmeasured prognostic patient characteristics randomly between groups, hence reducing bias, so that the difference in outcome can be interpreted as an effect of treatment. Baseline tables show whether randomization has led to comparable study

TABLE 4 Relationship of individual baseline imbalances with efficacy and risk of EPS and mortality

Imbalance between atypical antipsychotic and placebo group	Efficacy Change in SMD[a] (95% CI)	EPS Change in OR[a] (95% CI)	Mortality Change in OR[a] (95% CI)
Age in years, mean	−0.046 (−0.123; 0.030)	−0.0 (−0.7; 0.7)	0.4 (−0.9; 1.6)
Male gender, %	−0.027 (−0.048; −0.006)	−0.1 (−0.3; 0.1)	−0.2 (−0.5; 0.1)
Non-White race, %	−0.013 (−0.049; 0.023)	−0.4 (−0.8; −0.1)	−0.1 (−0.6; 0.4)
Vascular or mixed dementia, %[b]	0.015 (−0.011; 0.040)	−0.1 (−0.3; 0.4)	0.2 (−0.4; 0.7)
MMSE, mean[b]	−0.116 (−0.299; 0.067)	0.9 (−2.7; 4.5)	−1.0 (−4.0; 2.1)
NPS, standardized mean[b]	0.109 (−2.093; 2.311)	−7.9 (−30.0; 14.2)	2.7 (−9.4; 14.7)

Note. CI: confidence interval; EPS: extrapyramidal symptoms; MMSE: mini-mental state examination; NPS: neuropsychiatric symptoms; OR: odds ratio; SMD: standardized mean differences.

[a]per unit increase in the baseline difference.

[b]results based on less than 10 trials.

TABLE 5 Pooled efficacy, risk of EPS and risk of mortality for trials with reported and missing baseline information[a]

	Efficacy		EPS		Mortality	
Baseline characteristic	**SMD (95% CI)**		**OR (95% CI)**		**OR (95% CI)**	
Age in years, mean	Reported (12) Missing (2)	−0.102 (−0.173; −0.031) −0.243 (−0.390; −0.095)	Reported (11) Missing (1)	1.7 (1.3; 2.2) 1.6 (0.9; 2.7)	Reported (14) Missing (6)	1.7 (1.1; 2.6) 1.5 (0.6; 3.4)
Male gender, %	Reported/NA (11) Missing (3)	−0.100 (−0.172; −0.028) −0.243 (−0.387; −0.100)	Reported/NA (10) Missing (2)	1.8 (1.3; 2.3) 1.3 (0.8; 2.2)	Reported/NA (14) Missing (6)	1.7 (1.1; 2.6) 1.5 (0.6; 3.3)
Non-White race, %	Reported/NA (10) Missing (4)	−0.107 (−0.176; −0.038) −0.260 (−0.430; −0.090)	Reported/NA (9) Missing (3)	1.8 (1.3; 2.3) 1.4 (0.8; 2.2)	Reported/NA (12) Missing (8)	1.6 (1.1; 2.5) 1.8 (0.8; 4.4)
Vascular or mixed dementia, %[b]	Reported/NA (13) Missing (1)	−0.107 (−0.174; −0.040) −0.383 (−0.611; −0.155)	Reported/NA (11) Missing (1)	1.7 (1.3; 2.2) 1.6 (0.9; 2.7)	Reported/NA (19) Missing (1)	1.7 (1.1; 2.5) 1.5 (0.4; 5.4)
MMSE, mean	Reported (6) Missing (8)	−0.079 (−0.174; 0.016) −0.171 (−0.257; −0.084)	Reported (5) Missing (7)	2.0 (1.4; 2.8) 1.3 (0.9; 1.9)	Reported (7) Missing (13)	1.6 (0.9; 2.9) 1.7 (1.0; 2.8)
NPS, standardized mean	Reported/NA (3) Missing (11)	−0.042 (−0.181; 0.097) −0.152 (−0.225; −0.080)	Reported/NA (3) Missing (9)	2.4 (1.3; 4.5) 1.5 (1.2; 2.0)	Reported/NA (4) Missing (16)	1.2 (0.5; 2.6) 1.8 (1.2; 2.8)

Note. CI: confidence interval; EPS: extrapyramidal symptoms; MMSE: mini-mental state examination; OR: odds ratio; SMD: standardized mean differences; NA: not applicable, because some trials were performed in men only, in White persons only, in patients with Alzheimer disease only, and in patients without NPS at baseline.

[a]only trials without an extra haloperidol group.

[b]analyses performed with random effects models.

groups at the start of individual trials. Random fluctuations will still occur, but in general, the larger the sample size of an individual trial and the larger the number of trials in a review, the smaller the baseline imbalances can expected to be.

This is one of few studies that used objective measures to address risk of bias because of baseline imbalances in trials. Assessments of randomization are usually limited to the procedures, and these assessments can vary widely (Savovic et al., 2014).For example, using the Cochrane assessment tool, we found that 22 trials had an unclear random sequence generation and 22 trials an unclear concealment of allocation. In contrast, a Cochrane review reported that only four trials had unclear concealment of allocation (Ballard et al., 2006). Yet another review found that 100% of trials scored "high quality" on the Jadad and Van Tulder scale, and 90% on the Brown scale (Ma et al., 2014). We compared the true with expected group size difference and found that the distribution of differences was substantially smaller than could be expected by chance: 76% instead of 50% of the differences fell inside the 50% prediction interval.

4.2 | Baseline imbalances

CONSORT requires trial articles to present baseline tables for all randomized patients. We found four published trials that did not present a baseline table at all. Only a limited number of trials reported the six baseline characteristics we studied. Other characteristics that are likely to predict efficacy or adverse events, such as comorbidity and medication use, were also missing in many articles. Baseline information might not have been missing at random either. In our study, we found that trials with missing information had a more favorable pooled efficacy and risk of EPS than trials that provided the baseline information for each of the six characteristics. Selective reporting is a common problem in the medical scientific literature (Higgins et al., 2014), and missing information on prognostic baseline characteristics might be another example.

In the articles with baseline information, most of the imbalances seemed small but some were large and obviously clinically relevant. For example, in one study 30% of the participants receiving risperidone had vascular/mixed dementia versus 41% of the placebo group (De Deyn et al., 1999). The baseline imbalances that we investigated were not accounted for in the analyses of all but two trials (Kurlan et al., 2007; Schneider, Tariot, et al., 2006).

Our next step was to pool the baseline differences and assess heterogeneity, a method recently developed to quantify baseline differences (Clark et al., 2014, 2015). None of the pooled baseline differences we studied were statistically significant from zero. Some baseline differences showed considerable heterogeneity: The difference in mean age in trials with a haloperidol group (70%) and that in severity of NPS in trials without haloperidol group (27%). Heterogeneity for three characteristics was slight (between 5% and 11%). Perhaps, this amount of heterogeneity in baseline imbalances could be considered substantial as well, given that minimal heterogeneity is expected with an appropriate randomization design and conduct.

To quantify baseline imbalances, we also studied whether a positive or negative direction was more common. We found that numerically more trials reporting a higher mean age, higher percentage of men and lower severity of NPS, the primary focus of treatment in the trials, in the atypical antipsychotics group than in the placebo group. Others have suggested that imbalances in age and the primary outcome at baseline could be a good start when studying baseline imbalances.[2] We would like to add baseline imbalance in sex, and also differentiate between trials with and without a treatment arm with the old (patent free) competitor drug.

4.3 | Clinical outcomes

After assessing the presence of baseline imbalances, we investigated whether they might have affected the clinical outcomes of the trials. We found that higher mean age, higher percentage of men, and higher percentage of non-White persons at baseline in the antipsychotic than

the placebo group was associated with higher efficacy. For the baseline imbalance in sex, this was a statistically significant effect. Higher mean age, higher percentage of men, and higher percentage of non-White persons at baseline was also associated with a lower risk of EPS. For the baseline imbalance in race, this was a statistically significant effect. The effect of the baseline differences on risk of mortality was not so consistent but this was not a targeted outcome either. To our knowledge, there are no other studies that used this approach.

In addition, we found a consistent pattern of studies with missing baseline information having more favorable efficacy results and a lower risk of EPS on average. Naturally, the same studies with missing information having been pooled for each of the six baseline characteristics might partly underlie this finding. Again, the pattern was not consistent for the risk of mortality.

4.4 | Strengths and limitations

This is one of few studies that quantified baseline imbalances in trials. In addition, we performed an extensive literature search to identify unpublished studies. We hypothesized that baseline imbalances were related to outcomes, and hence the imbalance might depend on the publication status of a study. We used FDA and EMA databases among other literature sources (Schroll & Bero, 2015). The result was that we found six unpublished trials in addition to those included in previous meta-analyses (Ma et al., 2014; Schneider et al., 2005). As these were small studies and some did not report all outcomes, efficacy, risk of EPS, and risk of mortality for atypical antipsychotics versus placebo were not substantially different from those published before (Ma et al., 2014; Schneider et al., 2005).

A limitation of our study is that our analyses depended on the amount of baseline information provided in the articles. Information on type of dementia, MMSE, and severity of NPS was often lacking. Power of our study might have been insufficient to detect relevant baseline imbalances and associations of these baseline imbalances with clinical outcomes.

5 | CONCLUSION

Despite randomization, placebo-controlled trials of atypical antipsychotics in dementia show heterogeneous baseline imbalances. Baseline imbalances that were not taken into account might have mistakenly led to an overestimated efficacy and underestimated risk of EPS. Our findings underscore the need for adequate randomization procedures, and reporting of baseline characteristics for all randomized patients per treatment group. In addition, baseline imbalances need to be assessed objectively as part of systematic reviews.

DECLARATION OF INTEREST STATEMENT

None.

FUNDING INFORMATION

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

ACKNOWLEDGEMENTS

We would like to acknowledge the help of Professor De Deyn who provided additional data of two trials he participated in. In addition, we requested the individual patient data of the trials to assess the effect of baseline imbalances on pooled outcomes by adjusting for them in each trial before pooling the results. We decided not to go ahead with this study, because the majority of data were not provided. We would like to thank Dr. Paleacu and the National Institute of Mental Health for providing the data of their trials (Paleacu et al., 2008; Schneider, Tariot, et al., 2006).

AUTHOR CONTRIBUTIONS

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

ORCID

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

REFERENCES

Allain, H., Dautzenberg, P., Maurer, K., Schuck, S., Bonhomme, D., & Gerard, D. (2000). Double blind study of tiapride versus haloperidol and placebo in agitation and aggressiveness in elderly patients with cognitive impairment. *Psychopharmacology*, *148*, 361–366. https://doi.org/10.1007/s002130050064

Ballard, C., Margallo-Lana, M., Juszczak, E., Douglas, S., Swann, A., Thomas, A., … Jacoby, R. (2005). Quetiapine and rivastigmine and cognitive decline in Alzheimer's disease: Randomised double blind placebo controlled trial. *BMJ*, *330*(7496), 874. https://doi.org/10.1136/bmj.38369.459988.8F

Ballard, C., Waite, J., & Birks, J. (2006). Atypical antipsychotics for aggression and psychosis in Alzheimer's disease (Review). *The Cochrane Library*, *1*, 1–135.

Brodaty, H., Ames, D., Snowdon, J., Woodward, M., Kirwan, J., Ch, B., … Grossman, F. (2003). A randomized placebo-controlled trial of risperidone for the treatment of aggression, agitation, and psychosis of dementia. *The Journal of Clinical Psychiatry*, (February), *64*, 134–143. https://doi.org/10.4088/JCP.v64n0205

Carson, S., McDonagh, M., & Peterson, K. (2006). A systematic review of the efficacy and safety of atypical antipsychotics in patients with psychological and behavioral symptoms of dementia. *Journal of the American Geriatrics Society*, *54*(2), 354–361. https://doi.org/10.1111/j.1532-5415.2005.00566.x

Cheung, G., & Stapelberg, J. (2011). Quetiapine for the treatment of behavioural and psychological symptoms of dementia (BPSD): A meta-analysis of randomised placebo-controlled trials. *The New Zealand Medical Journal*, *124*(1336), 39–50.

Clark, L., Fairhurst, C., Cook, E., & Torgerson, D. (2015). Important outcome predictors showed greater baseline heterogeneity than age in two systematic reviews. *Journal of Clinical Epidemiology*, *68*(2), 175–181. https://doi.org/10.1016/j.jclinepi.2014.09.023

Clark, L., Fairhurst, C., Hewitt, C., Birks, Y., Brabyn, S., Cockayne, S., … Torgerson, D. (2014). A methodological review of recent meta-analyses has found significant heterogeneity in age between randomized groups. *Journal of Clinical Epidemiology*, *67*(9), 1016–1024. https://doi.org/10.1016/j.jclinepi.2014.04.007

Corbett, M., Higgins, J., & Woolacott, N. (2014). Assessing baseline imbalance in randomised trials: Implications for the Cochrane risk of bias

tool. *Research Synthesis Methods*, *5*(1), 79–85. https://doi.org/10.1002/jrsm.1090

De Deyn, P., Carrasco, M., Deberdt, W., Jeandel, C., Hay, D., Feldman, P., … Breier, A. (2004). Olanzapine versus placebo in the treatment of psychosis with or without associated behavioral disturbances in patients with Alzheimer's disease. *International Journal of Geriatric Psychiatry*, *19*(2), 115–126. https://doi.org/10.1002/gps.1032

De Deyn, P., Jeste, D., Swanink, R., Kostic, D., & Breder, C. (2005). Aripiprazole for the treatment of psychosis in patients with Alzheimer's disease. *Journal of Clinical Psychopharmacology*, *25*(5), 463–467. Retrieved from http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00004714-200510000-00010. https://doi.org/10.1097/01.jcp.0000178415.22309.8f

De Deyn, P., Rabheru, K., & Rasmussen, A. (1999). A randomized trial of risperidone, placebo, and haloperidol for behavioral symptoms of dementia. *Neurology*, *53*, 945–956.

Deberdt, W., Dysken, M., Rappaport, S., Feldman, P., Young, C., & Hay, D. (2005). Comparison of olanzapine and risperidone in the treatment of psychosis and associated behavioral disturbances in patients with dementia. *The American Journal of Geriatric Psychiatry*, *13*(August), 722–730. https://doi.org/10.1097/00019442-200508000-00012

Deeks, J., Higgins, J., & Altman, D. (2011). Chapter 9: Analysing data and undertaking meta-analyses. In *Cochrane handbook for systematic reviews of interventions (Version 5., Vol. Version 5.)*. The Cochrane Collaboration.

Ebell, M. (2013). Methodological concerns about studies on oseltamivir for flu. *BMJ*, *347*, f7148. https://doi.org/10.1136/bmj.f7148

Egger, M., Juni, P., Bartlett, C., Holenstein, F., & Sterne, J. (2003). How important are comprehensive literature searches and the assessement of trial quality in systematic reviews? Emperical study. *Health Technology Assessment*, *7*(1).

Fitzpatrick, A., Kuller, L., Lopez, O., Kawas, C., & Jagust, W. (2005). Survival following dementia onset: Alzheimer's disease and vascular dementia. *Journal of the Neurological Sciences*, *229–230*, 43–49.

Gambassi, G., Landi, F., Lapane, K. L., Sgadari, A., Mor, V., & Bernabei, R. (1999). Predictors of mortality in patients with Alzheimer's disease living in nursing homes. *Journal of Neurology, Neurosurgery and Psychiatry*, *67*, 59–65. https://doi.org/10.1136/jnnp.67.1.59

Garcia-Ptacek, S., Farahmand, B., Kåreholt, I., Religa, D., Cuadrado, M., & Eriksdotter, M. (2014). Mortality risk after dementia diagnosis by dementia type and underlying factors: A cohort of 15,209 patients based on the Swedish Dementia Registry. *Journal of Alzheimer's Disease*, *41*, 467–477. https://doi.org/10.3233/JAD-131856

Haupt, M., Cruz-Jentoft, A., & Jeste, D. (2006). Mortality in elderly dementia patients treated with risperidone. *Journal of Clinical Psychopharmacology*, *26*(6), 566–570. https://doi.org/10.1097/01.jcp.0000239796.21826.39

Hernán, M., Hernández-Díaz, S., & Robins, J. (2004). A structural approach to selection bias. *Epidemiology*, *15*(5), 615–625. https://doi.org/10.1097/01.ede.0000135174.63482.43

Herz, L., Volicer, L., Frankenburg, F., Colon, S., & Kittur, S. (2002). A 6-week, double-blind comparison of olanzapine, risperidone, and placebo for behavioral disturbance in Alzheimer's disease. *The Journal of Clinical Psychiatry*, *63*(11), 1065.

Higgins, J., Altman, D., & Sterne, J. (2014). Chapter 8: Assessing risk of bias in included studies. In *Cochrane handbook*. The Cochrane Collaboration. Retrieved from http://handbook.cochrane.org/chapter_8/8_assessing_risk_of_bias_in_included_studies.htm

Hróbjartsson, A., Boutron, I., Turner, L., Altman, D., & Moher, D. (2013). Assessing risk of bias in randomised clinical trials included in Cochrane reviews: The why is easy, the how is a challenge [editorial]. *Cochrane Database of Systematic Reviews*, *4*. https://doi.org/10.1002/14651858.ED000058

Hulshof, T., Zuidema, S., Ostelo, R., & Luijendijk, H. (2015). The Mortality Risk of Conventional Antipsychotics in Elderly Patients: A Systematic Review and Meta-analysis of Randomized Placebo-Controlled Trials. *Journal of the American Medical Directors Association*, *16*(10), 817–824. https://doi.org/10.1016/j.jamda.2015.03.015

ILO522. (2002). A prospective, randomized, double-blind, placebo-controlled, flexible-dose, parallel-group, single-center study to evaluate the safety, tolerability and efficacy of iloperidone (0.5 to 6.0 mg/d given b.i.d.) compared with placebo in treating psychotic and.

Katz, I., Jeste, D., Mintzer, J., Clyde, C., Napolitano, J., & Brecher, M. (1999). Comparison of risperidone and placebo for psychosis and behavioral disturbances associated with dementia: A randomized, double-blind trial. *The Journal of Clinical Psychiatry*, *60*(February), 107–115. https://doi.org/10.4088/JCP.v60n0207

Kennedy, J., Deberdt, W., Siegal, A., Micca, J., Degenhardt, E., Ahl, J., … Baker, R. (2005). Olanzapine does not enhance cognition in non-agitated and non-psychotic patients with mild to moderate Alzheimer's dementia. *International Journal of Geriatric Psychiatry*, *20*(11), 1020–1027. https://doi.org/10.1002/gps.1397

Kurlan, R., Cummings, J., Raman, R., & Thal, L. (2007). Quetiapine for agitation or psychosis in patients with dementia and parkinsonism. *Neurology*, *68*(17), 1356–1363. https://doi.org/10.1212/01.wnl.0000260060.60870.89

Lee, P., Gill, S., Freedman, M., Bronskill, S., Hillmer, M., & Rochon, P. (2004). Atypical antipsychotic drugs in the treatment of behavioural and psychological symptoms of dementia: Systematic review. *BMJ*, *329*(June), 1–5.

Lonergan, E., Luxenberg, J., Colford, J., & Birks, J. (2010). Haloperidol for agitation in dementia (review). *The Cochrane Library*, *12*, 1–42.

Luijendijk, H., & Hulshof, T. (2015). Baseline differences in the SAVOR trial. *Diabetes, Obesity & Metabolism*, *17*(12), 1202. https://doi.org/10.1111/dom.12505

Ma, H., Huang, Y., Cong, Z., Wang, Y., Jiang, W., Gao, S., & Zhu, G. (2014). The efficacy and safety of atypical antipsychotics for the treatment of dementia: A meta-analysis of randomized placebo-controlled trials. *Journal of Alzheimer's Disease*, *42*, 915–937. https://doi.org/10.3233/JAD-140579

Maher, A., Maglione, M., Bagley, S., Suttorp, M., Hu, J., Ewing, B., … Shekelle, P. G. (2011). Efficacy and comparative effectiveness of atypical antipsychotic medications for off-label uses in adults. *Journal of the American Medical Association*, *306*(12), 1359–1370. https://doi.org/10.1001/jama.2011.1360

Mintzer, J., Greenspan, A., Caers, I., Hove, I., Kushner, S., Weiner, M., … Schneider, L. (2006). Risperidone in the treatment of psychosis of Alzheimer disease: Results from a prospective clinical trial. *The American Journal of Geriatric Psychiatry*, (March), *14*, 280–291. https://doi.org/10.1097/01.JGP.0000194643.63245.8c

Mintzer, J., Tune, L., Breder, C., Swanink, M., & Marcus, R. (2007). Aripiprazole for the treatment of psychoses in institutionalized patients with Alzheimer dementia: A multicenter, randomized, of three fixed doses. *The American Journal of Geriatric Psychiatry*, (November), 918–931.

Mitchell, S., Kiely, D., Hamel, M., Park, P., Morris, J., & Fries, B. (2004). Residents with advanced dementia. *JAMA*, *291*(22), 2734–2740. https://doi.org/10.1001/jama.291.22.2734

Paleacu, D., Barak, Y., Mirecky, I., & Mazeh, D. (2008). Quetiapine treatment for behavioural and psychological symptoms of dementia in Alzheimer's disease patients: A 6-week, double-blind, placebo-controlled study. *International Journal of Geriatric Psychiatry*, *23*(September 2007), 393–400. https://doi.org/10.1002/gps.1892

R Core Team (2013). *R: A language and environment for statistical computing*. Vienna, Austria.

Riley, R., Kauser, I., Bland, M., Thijs, L., Staessen, J., Wang, J., … Deeks, J. (2013). Meta-analysis of randomised trials with a continuous outcome according to baseline imbalance and availability of individual participant data. *Statistics in Medicine*, *32*(16), 2747–2766. https://doi.org/10.1002/sim.5726

RIS-BEL-14. (1997). Unpublished.

RIS-INT-83. (1997). Unpublished.

Rountree, S., Chan, W., Pavlik, V., Darby, E., & Doody, R. (2012). Factors that influence survival in a probable Alzheimer disease cohort. *Alzheimer's Research & Therapy*, *4*(3), 119. https://doi.org/10.1186/alzrt119

Satterlee, W. (1995). Olanzapine versus placebo in the treatment of patients with psychosis associated with dementia (F1D-MC-HGAO).

Savovic, J., Weeks, L., Sterne, J., Turner, L., Altman, D., Moher, D., & Higgins, J. (2014). Evaluation of the cochrane collaboration's tool for assessing the risk of bias in randomized trials: focus groups, online survey, proposed recommendations and their implementation. *Systematic Reviews*, *3*(1), 1–12. Retrieved from Systematic Reviews

Schneider, L., Dagerman, K., & Insel, P. (2005). Risk of death with atypical antipsychotic drug treatment for dementia. *Journal of the American Medical Association*, *294*(15), 1934–1943. https://doi.org/10.1001/jama.294.15.1934

Schneider, L., Dagerman, K., & Insel, P. (2006). Efficacy and adverse effects of atypical antipsychotics for dementia: Meta-analysis of randomized, placebo-controlled trials. *American Journal of Geriatric Psychiatry*, *14*(March), 191–210. https://doi.org/10.1097/01.JGP.0000200589.01396.6d

Schneider, L., Tariot, P., Dagerman, K., Davis, S., Hsiao, J., Ismail, S., & Lebowitz, B. (2006). Effectiveness of atypical antipsychotic drugs in patients with alzheimer's disease. *The New England Journal of Medicine*, *355*(15), 1525–1538. https://doi.org/10.1056/NEJMoa061240

Schroll, J., & Bero, L. (2015). Regulatory agencies hold the key to improving cochrane reviews of drugs [editorial]. *Cochrane Database of Systematic Reviews*, *4*. https://doi.org/10.1002/14651858.ED000098

Schulz, K., Altman, D., & Moher, D. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *Journal of Clinical Epidemiology*, *63*(8), 1–6. https://doi.org/10.1016/j.ijsu.2010.09.006

Schulz, K., Chalmers, I., Grimes, D., & Altman, D. (1994). Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *JAMA*, *272*(2), 125–128. https://doi.org/10.1001/jama.1994.03520020051014

Schulz, K., Chalmers, I., Hayes, R., & Altman, D. (1995). Empirical evidence of bias. dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*, *273*, 408–412. https://doi.org/10.1001/jama.1995.03520290060030

Seitz, D., Gill, S., Herrmann, N., Brisbin, S., Rapoport, M., Rines, J., … Conn, D. (2013). Pharmacological treatments for neuropsychiatric symptoms of dementia in long-term care: A systematic review. *International Psychogeriatrics*, *25*(2), 185–203. https://doi.org/10.1017/S1041610212001627

StataCorp (2013). *Stata statistical software: Release 13*. College Station, TX: StataCorp LP.

Street, J., Clark, S., Gannon, K., Cummings, J., & Bymaster, F. (2000). Olanzapine treatment of psychotic and behavioral symptoms in patients with Alzheimer disease in nursing care facilities. *Archives of General Psychiatry*, *57*, 968–976. https://doi.org/10.1001/archpsyc.57.10.968

Streim, J., Porsteinsson, A., Breder, C., Swanink, R., & Marcus, R. (2008). Placebo-controlled study of aripiprazole for the treatment of psychosis in nursing home patients with Alzheimer disease. *The American Journal of Geriatric Psychiatry*, (July), 537–550.

Sweeting, M., Sutton, A., & Lambert, P. (2004). What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Statistics in Medicine*, *23*(9), 1351–1375. https://doi.org/10.1002/sim.1761

Tariot, P., Schneider, L., Katz, I., Mintzer, J., Street, J., Copenhaver, M., & Williams-Hughes, C. (2006). Quetiapine treatment of psychosis associated with dementia: A double-blind, randomized, placebo-controlled clinical trial. *The American Journal of Geriatric Psychiatry*, *14*(9), 767–776. https://doi.org/10.1097/01.JGP.0000196628.12010.35

Trowman, R., Dumville, J., Torgerson, D., & Cranny, G. (2007). The impact of trial baseline imbalances should be considered in systematic reviews: A methodological case study. *Journal of Clinical Epidemiology*, *60*(12), 1229–1233. https://doi.org/10.1016/j.jclinepi.2007.03.014

US Food and Drug Administration. (2013). Atypical Antipsychotic Drugs Information. Retrieved from http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm094303.htm

World Health Organization. (2013). WHO Collaborating Centre for Drug Statistics Methodology. Retrieved from http://www.whocc.no/atc_ddd_index/?code=N05A

Zhong, K., Tariot, P., Mintzer, J., Minkwitz, M., & Devine, N. (2007). Quetiapine to treat agitation in sementia: A randomized, double-blind, placebo-controlled study. *Current Alzheimer Research*, *4*(1), 81–93. https://doi.org/10.2174/156720507779939805

ZIP-128-105. (1993). Third party blind, evaluation of the clinical usefulness of low dose ziprasidone for the symptomatic treatment of dementia in institutionalized patients.

SUPPORTING INFORMATION

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

How to cite this article: Hulshof TA, Zuidema SU, van Meer PJK, Gispen-de Wied CC, Luijendijk HJ. Baseline imbalances and clinical outcomes of atypical antipsychotics in dementia: A meta-epidemiological study of randomized trials. *Int J Methods Psychiatr Res*. 2019;28:e1757. https://doi.org/10.1002/mpr.1757