

Efficacy of Atypical Antipsychotics to Treat Behavioral and Psychological Symptoms of Dementia in Nursing Home Residents: A Systematic Review of the Evidence from Randomized Controlled Trials

Keywords: Atypical antipsychotics; Behavioral and psychological symptoms of dementia; Dementia; Nursing home

Abstract

Background: Both behavioral and psychological symptoms of dementia and the use of atypical antipsychotics among nursing home residents with dementia are epidemic. However, the evidence from systematic reviews and randomized controlled trials for the benefits of using atypical antipsychotics in this population is conflicting and inconsistent. This study examines the evidence from randomized controlled trials for the use of atypical antipsychotics to treat nursing home residents with behavioral and psychological symptoms of dementia.

Methods: PubMed, Cochrane review, the National Clinical Guideline Clearinghouse, previously published systematic reviews, and a search of references were used to find eligible randomized controlled trials on use of atypical antipsychotics to treat NH residents with behavioral and psychological symptoms of dementia. Inclusion and exclusion criteria were pre-defined. The papers were independently reviewed by two investigators. Risk of bias, applicability, and heterogeneity were assessed based on previously published methods.

Results: Among 1469 citations, the 12 trials met the inclusion criteria. The variability of the diagnostic criteria and outcome measurements, as well as that of these verity of dementia and behavioral and psychological symptoms of dementia, prohibited meta-analysis of the selected 12 original studies. Among the total 4352 enrolled subjects across the 12 trials, age ranged from 77 to 84 years old. Seventy two percent of participants were female, and 64% of participants completed the trials. Mean duration of the 12 trials varied from six to 12 weeks. Four out of the 12 trials (33%) showed positive results. The BPSD reduction varied from 7% to 72%. Risk of bias included low concealment (58%) and high attrition rate (20-42%). The strict exclusion criteria and low recruitment fractions (69%) among the 12 trials reduced the applicability of the trials. Additionally, there was significant clinical heterogeneity and methodological diversity across the 12 trials.

Conclusions: There is limited and inconsistent evidence to demonstrate the efficacy of atypical antipsychotics in the treatment of behavioral and psychological symptoms of dementia in nursing home residents. In addition, there are concerns about risk of bias, applicability, clinical heterogeneity, and methodological diversity among the 12 trials. We are less confident in the intervention effects of atypical antipsychotics on BPSD, and therefore do not recommend routinely prescribing atypical antipsychotics.



***Address for Correspondence**

Huai Yong Cheng, MD, MPH, Division of General Medicine, Geriatrics, and Palliative Care, Department of Medicine, University of Virginia, PO Box 800901, Charlottesville, VA 22981, USA, Tel: 434-924-1685; Fax: 434-977-0581; E-mail: hyc9j@hscmail.mcc.virginia.edu

Copyright: © 2015 Haight TN, et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Submission: 05 February 2015

Accepted: 11 May 2015

Published: 16 May 2015

Abbreviations

BEHAVE-AD: Behavioral Pathology in Alzheimer's disease; BPRS: Brief Psychiatric Rating Scale; BPSD: Dementia and Behavioral and Psychological Symptoms of Dementia; CMAI: Cohen-Mansfield Agitation Inventory; CMS: Center for Medicare and Medicaid; CONSORT: Consolidated Standards of Reporting Trials; DSM: Diagnostic and Statistical Manual of Mental Disorders; FAST: Functional Assessment Staging Test; FDA: Food and Drug Administration; NH: Nursing Home; IRB: Institutional Review Board; MMSE: Mini-Mental State Examination; NINCDS-ADRDA: National Institute of Neurological and Communicative Diseases and Stroke-Alzheimer's Disease and Related Disorders Association; NPS: Neuropsychiatric Symptoms; NPI: Neuropsychiatric Inventory; NPI-NH: Neuropsychiatric Inventory-Nursing Home; OIG: Office of the Inspector General; PANSS-EC: Positive and Negative Syndrome Scale-Excitement Component; RCTs: Randomized Controlled Trials

Introduction

Behavioral and psychological symptoms of dementia (BPSD) are defined as "symptoms of disturbed perception, thought content, mood or behavior that frequently occur in patients with dementia" according to the International Psychogeriatric Association [1]. The International Psychogeriatric Association also states that "behavioral symptoms are usually identified on the basis of observation of the patient, including physical aggression, screaming, restlessness, agitation, wandering, culturally inappropriate behaviors, sexual disinhibition, hoarding, cursing and shadowing. Psychological symptoms are usually and mainly assessed on the basis of interviews with patients and relatives. These symptoms include anxiety, depressive mood, hallucinations and delusions and psychosis" [1]. Because 23 BPSD scales and checklists have been published in the literature [1], we use BPSD as a term to cover behavioral and psychological symptoms or NPS in this review. BPSD among nursing home (NH) residents are common and eventually occur to every NH resident [1-4] and potentially cause significant distress and burden to both the individuals with dementia and their caregivers [5-10]. Consequently, antipsychotics,

mainly atypical agents, have been widely used in the NH setting [11-17], resulting in concerns for adverse effects and regulations from several government agencies including the Office of the Inspector General (OIG), the Food and Drug Administration (FDA), and the Center for Medicare and Medicaid (CMS) [18-20]. Atypical antipsychotic use is not only associated with high cost, but also with multiple adverse effects including increased mortality, falls, strokes, and myocardial infarction [1,18-30]. Additionally, the evidence from systematic reviews and randomized controlled trials (RCTs) for the benefit of atypical antipsychotics use in this vulnerable population is inconsistent and conflicting [31-43]. Ten systematic reviews examining BPSD suggest that atypical antipsychotics are effective [31-40]. In contrast, three systematic reviews on BPSD suggest that atypical antipsychotics are ineffective [41-43]. Most importantly, all of these systematic reviews [31-40,42,43] except one [41] do not focus on NH residents with BPSD. One systematic review without a meta-analysis included 15 studies examining atypical and typical antipsychotics to treat BPSD in NH residents and found atypical antipsychotics to be ineffective for treating BPSD in NH residents [41]. In the absence of consistent good evidence, many general review articles have recommended the use of atypical antipsychotics for BPSD in older patients [44-55], though the guidelines provide inconsistent and conflict recommendations [56-60]. Importantly, a systematic review focusing on treating BPSD in NH residents with atypical antipsychotics has not been previously reported. Therefore, we decided to perform this systematic review. The PICOS (patient, intervention, comparison, outcome, study design) question [61,62] of this systematic review is used to examine whether the use of atypical antipsychotics reduces BPSD in NH residents compared to placebo or usual care based on the evidence from RCTs. Specifically, the objectives of this systematic review are to address the following four questions: 1) Can atypical antipsychotics reduce BPSD among NH residents? 2) Are the RCTs examining atypical antipsychotics to treat BPSD in NH residents valid? i.e., is the internal validity good? 3) What is the role of the placebo effect of atypical antipsychotics in the treatment of BPSD for NH residents? 4) Are the results of RCTs applicable to the population in the real world? i.e., is the external validity good? We also provide a list of BPSD measurements and scales, which were not reported in all previous systematic reviews [31-43].

Methods

General approach

In performing this systematic review, we followed the standards of the Cochrane Handbook for Systematic Reviews of Intervention [61] and Preferred Reporting Items for Systematic Reviews and Meta-analyses [62] with the following modifications. We did not use scale scores such as Jahad score to assess the quality of RCTs in our previous study [63]. Instead, we adopted a content evaluation method to assess the internal validity and risk of bias of the trials [61-64]. Assessment of the internal and external validity of the RCTs followed the Consolidated Standards of Reporting Trials (CONSORT) [64] and methods from our previous study [63] and others [65].

Protocol and registration

The first two authors had several meetings to discuss pre-defined research questions, eligible criteria, search strategy, data collection,

assessment of internal validity and external validity, analytic approach, and summary of results during the planning phase. However, we did not write a full protocol and complete the online registration.

Eligibility criteria

We defined the eligibility criteria of this systematic review as following based on our PICOS research question: 1) Participants: NH residents with BPSD. The rationale is that BPSD in NH patients is epidemic and a significant problem [2-4] and atypical antipsychotics are widely used in this vulnerable population [11-17]. Additionally, the evidence from previous systematic reviews based on RCTs is inconsistent and conflicting [41-43]; 2) Intervention: atypical antipsychotics as the intervention; 3) Comparison: placebo or usual care; 4) Outcome measures: BPSD used as the primary outcome explicitly or implicitly; 5) Study design: parallel RCTs that lasted at least six weeks. RCTs are well-known to be the gold standard for testing the efficacy of a given intervention. Full papers (not the abstract) were published in English. Non-randomized trials, trials without a control, crossover or head-to-head designs, open-label trials, duplicate reporting of the same trial and reporting in the format of an abstract were excluded.

Information sources

The Medline database (PubMed) was used. This search was conducted from the inception to October 19, 2013. Additional research included Cochrane review, the National Guideline Clearinghouse database, previously published systematic reviews [31-43], and a hand search for references. Unpublished studies were not examined.

Search strategy

Starting with PubMed, we searched for relevant articles using MeSH terms, key words, and text words. The search was conducted in the following six steps: Step 1 for the RCTs Domain: using the following term “randomized controlled trial or controlled clinical trial or placebo or randomly or trial or group or drug therapy or randomized” to retrieve all RCT-relevant citations; Step 2 for Atypical Antipsychotics Domain: using the following term “ziprasidone or quetiapine or olanzapine or aripiprazole or risperidone or clozapine or amisulpride or sertindole or zotepine or atypical antipsychotics” to retrieve all citations relevant to atypical antipsychotics; Step 3 for the BPSD Domain: using the following term “neuropsychiatric symptoms or psychiatric disorders or behavior symptoms or behavioral symptoms dementia or psychological symptoms or behavioral and psychological symptoms of dementia or behavioral disturbances or agitation or aggression or delusion or hallucination or disinhibition or wandering or irritability or delirium or psychosis or euphoria or anxiety or apathy or aberrant motor behavior or night time behavior disturbances or appetite and eating abnormalities” to retrieve all BPSD-related citations; Step 4 for the Dementia Domain: using the following term “dementia or delirium or amnestic dementia or cognitive disorders or vascular dementia or AIDS dementia complex or multi-infarct dementia or front temporal dementia or Alzheimer vaccine or Alzheimer’s disease or cognition disorders or cognitive impairment or mild cognitive impairment” to retrieve all dementia-related citations. Step 5 for NH domain: nursing homes, assisted living facilities, homes for aged, assisted living, residential

facilities, housing for the elderly, skilled nursing facilities, long term care, skilled nursing, long term care facilities; Step 6 for combining all citations from Step 1 to 5: Using “AND” to combine RCTs Domain, Atypical Antipsychotics Domain, BPSD Domain, Dementia Domain, and NH domain together and get all related citations for screening eligible papers. The search strategy was guided by an experienced medical librarian from our institution. The search terms and search domains for this systematic review were saved in PubMed (available as requested).

To avoid any missing studies, we also searched Cochrane review, the National Guideline Clearinghouse database, and previously published systematic reviews [31-43], and conducted a hand search for references as well.

Study selection process

All possible citations were retrieved via PubMed. One investigator (HYC) screened and identified all citations for potentially eligible studies. Study selection included the following two steps: 1) Examine the title and abstract for possible inclusion for this study; 2) Review the original article for definite inclusion into this study (Figure 1).

Data extraction and collection process

We modified the data extraction sheet based on our previous study [63]. Studies that met the inclusion criteria were reviewed independently. The data sheet was completed independently by two investigators (HYC and TNH). The same two investigators met and compared the completed data sheets, and any discrepancies were resolved by face to face discussion. One author (TNH) entered the data into the Excel data set, which was further reviewed by another author (HYC) for accuracy. The data in Excel was imported to SPSS by one author (HYC) who also performed data cleaning, coding and analysis. The content in all tables was independently reviewed by all authors.

Data items

The data sheet included the first author name of the publication, year of the publication, the journal name, methods and scales of dementia and depression diagnostic criteria, demographics (age, gender, and race), measurements and scales of primary outcome measures, secondary outcomes, intervention and comparison, participant recruitment processes including the number of eligible and enrolled participants, drug interventions, trial duration, power calculations, randomization status, blinding status, intention-to-treat, outcome measures and statistical significance, sponsorship, Institutional Review Board (IRB) approval, and informed consent. Placebo effects from the intervention and comparison groups from the 12 trials were later added to the sheet and Table 5 by one investigator (HYC). Summary of exclusion criteria and adverse effects were added to Tables 4 and 5.

Risk of bias in individual studies

We assessed risk of bias by assessing the internal validity which is defined as whether the study results in true findings and minimizes systematic error. The internal validity of the trials was determined based on the quality assessment of content from the previously published methods as well as from a previous study of one of the authors (HYC), and included randomization, double blinding,

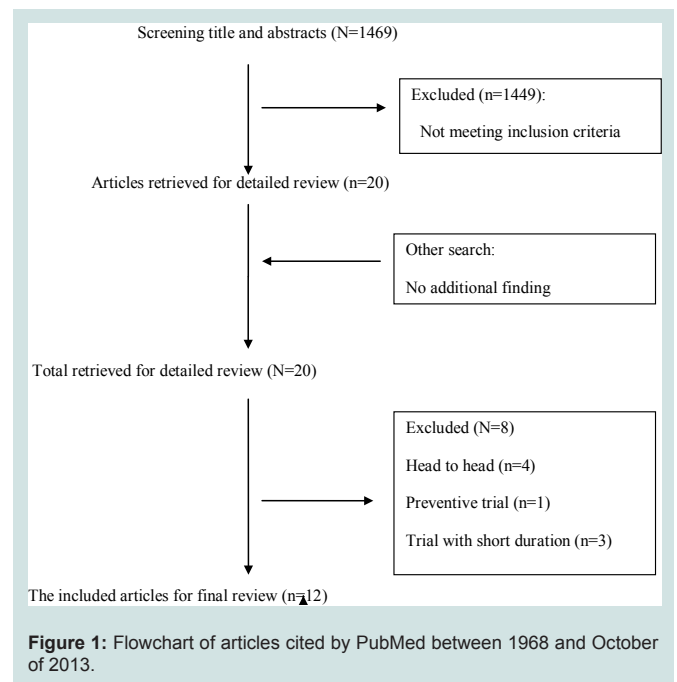


Figure 1: Flowchart of articles cited by PubMed between 1968 and October of 2013.

concealment, attrition, intention-to-treat analysis, and a power calculation [61-64]. Assessment of these items covered selection bias (concealment), performance bias (blinding), attrition bias (attrition), and detection bias (blinding) [61].

Applicability in individual studies

We assessed the applicability of individual studies by examining external validity i.e., whether the results of the studies in the research setting can be applied to the population in the real world. External validity was assessed by examining the recruitment process, i.e., the percentage of patients in the daily practice that was enrolled through the recruitment process and the exclusion of research subjects [61-65]. The eligible index was defined as the percentage of patients who were screened and met the inclusion criteria. The enrollment index was defined as the percentage of eligible patients who were enrolled in the trial. The recruitment index was defined as the percentage of patients who were screened and enrolled in the trial [63-65]. The exclusion criteria reported among the trials included certain medications, medical diseases, psychiatric and mental illness, and other reasons which are summarized in Table 4.

Heterogeneity among the selected trials

Heterogeneity is defined as any kind of variability among studies. It includes 1) clinical diversity (also called clinical heterogeneity), i.e., variability in the research participants, intervention, outcomes; 2) methodological diversity, i.e. variability in study design and risk of bias; and 3) statistical heterogeneity, i.e., variability in the intervention effects [61]. Our inclusion criteria cover clinical heterogeneity e.g., intervention, outcome measurements, and external validity and methodological diversity e.g., risk of bias, internal validity. We did not perform a meta-analysis. Therefore, we did not include an assessment of statistical heterogeneity.

Synthesis of results

We used percentage of BPSD reduction as the primary outcome

Table 1: Summary of trial profile in the 12 selected trials on treating behavioral psychiatric symptoms of dementia in nursing home residents with dementia.

Year of publication, Last name of first author	Country & setting	Sample size		Age (year)	Gender (% female)	Race (% white)	Dementia diagnostic criteria and severity of dementia in intervention and control groups	BPSD diagnosis criteria and severity of BPSD in intervention and control groups
		Target (screened) population	Enrolled					
2008 Streim [73]	U.S. NH or residential assisted living	330	256	83	76	90	DSM IV and MMSE (6-22) Intervention: moderate dementia (Mean MMSE=14) Control: moderate dementia (Mean MMSE=13)	Psychotic symptom of delusion or hallucination at least intermittently ≥1 month. A score of ≥6 on either delusion or hallucinations items of NPI-NH Intervention: mean NPI-NH psychosis score=10, mean CGI-S=4.4 Control: mean NPI-NH psychosis score=11, mean CGI-S=4.5
2007 Mintzer [74]	U.S. NH or residential assisted- living facilities	unknown	487	82	79	87	DSM IV and MMSE (6 and 22). Intervention and control group: moderate dementia (Mean MMSE=12)	Persistent or intermittent delusion or hallucination or both ≥1 month. A score of ≥6 on either delusion or hallucinations items of NPI-NH psychosis subscale score Intervention and control group: mean NPI-NH psychosis score=12, mean CGI-S=4.7
2007 Zhong [75]	U.S. NH or assisted living facilities	435	333	83	74	84	DSM IV (possible AD or vascular dementia) or NINCDS/ADRDA Intervention: severe dementia (Mean MMSE=5) Control: severe dementia (Mean MMSE=6)	Total score ≥14 on PANSS-EC and a score ≥4 on one of five PANSS-EC subscale Intervention: mean PANSS-EC total score=23, mean CGI-S=4.7 Control: mean PANSS-EC total score=23, mean CGI-S=4.8
2006 Mintzer [76]	U.S. NH or long-term care facilities	560	473	83	77	80	Not defined. MMSE was (5-23). Intervention: moderate dementia (mean MMSE=13) Control: moderate dementia (mean MMSE=13)	Psychosis of AD and ≥2 on any items of BEHAVE-AD psychosis subscale Intervention: mean BEHAVE-AD psychosis subscale=7, mean CGI-S=3.3 Control: mean BEHAVE-AD psychosis subscale=8, mean CGI-S=3.3
2006 Tariot [77]	U.S.	501	190	83	76	88	DSM IV (probable AD) or NINCDS-ADRA (possible AD) and MMSE≥5. Intervention: moderate dementia (mean MMSE=12) Control: moderate dementia (mean MMSE=13)	BPRS≥24, CGI-S≥4, frequency scores of ≥ 3 on at least one of two psychosis (delusion or hallucinations) of NPI-NH Intervention: mean BPRS total score=40, mean CGI-S=5 Control: mean BPRS total score=39, mean CGI-S=5
2005 Ballard [78]	England NH Care facilities	282	62	84	82 (exclude rivastigmine)	unknown	NINCDS-ADRDA (probable or possible AD) Intervention: severe dementia (mean FAST=6, mean SIB=59) Control: severe dementia (mean FAST=6, mean SIB=69)	CMAI>39, clinically significant agitation at least 6 weeks, and score≥4 on irritability or aberrant motor behavior scales of NPI for four week Intervention: mean CMAI=59 Control: mean CMAI=56
2003 Brodaty [79]	Australia NH	unknown	345	83	72	Unknown	DSM IV, FAST≥4, MMSE≤23 Intervention: severe dementia (mean MMSE=5) Control: severe dementia (mean MMSE=6)	A minimum aggressive score of CMAI: a score ≥4 on at least 1 aggressive items, or a score of 3 on at least 2 aggressive items, or a score of 2 on at least 3 aggressive items, or 2 aggressive items occurring at a frequency of 2 and 1 at a frequency of 3. Intervention: mean CMAI score=34, mean BEHAVE-AD=19 Control: mean CMAI score=33, mean BEHAVE-AD =19
2000 Street [80]	U.S. NH	288	206	83	61	93	NINCDS-ADRDA (probable or possible AD), MMSE<24 Intervention: severe dementia (mean MMSE=7) Control: severe dementia (mean MMSE=7)	scores≥3 on any of agitation/aggression, hallucinations or delusions items of NPI-NH Intervention: mean total score of NPI-NH=14 Control: mean total score of NPI-NH=15

1999 De Deyn [81]	Europe and Canada Institutionalized patients	unknown	229	81	58	99	DSM IV, FAST score ≥ 4 , MMSE ≤ 23 Intervention: severe dementia (mean MMSE=9) Control: severe dementia (mean MMSE=9)	>1 on BEHAVE-AD global rating, ≥ 8 on BEHAVE-AD total score. Intervention: mean BEHAVE-AD total score=16 Control: mean BEHAVE-AD total score=17
2005 Deberdt [82]	U.S. NH or assisted living centers or outpatient	unknown	494	79	65	84	DSM IV, NINCDS-ADRDA Intervention: moderate dementia (mean MMSE=14) Control: moderate dementia (mean MMSE=15)	Scores ≥ 6 (severity X frequency) on sum of hallucinations and delusions items of NPI or NPI-NH Intervention: mean NPI total score=12 Control: mean NPI total score=11
2004 De Deyn [83]	Europe, Australia, Israel, Lebanon, South Africa Long-term NH or continuing-care hospitals	unknown	652	77	75	99.7	DSM IV-TR (probable or possible AD), NINCDS-ADRDA, MMSE ≥ 5 . Intervention: mild to severe dementia (mean MMSE=14) Control: mild to severe dementia (mean MMSE=14)	Delusion or hallucinations (no cut off of NPI-NH psychosis total score) Intervention: mean NPI-NH psychosis total score=10, mean CGI-S=5 Control: mean NPI-NH psychosis total score=10, mean CGI-S=5
1999 Katz [84]	U.S. NH or chronic disease hospital	729	625	83	68	89	DSM IV, MMSE ≤ 23 , FAST score ≥ 4 Intervention: severe dementia (mean MMSE=6-8, 46% $\geq 7A$ on FAST) Control: severe dementia (mean MMSE=6, 49% $\geq 7A$ on FAST)	BEHAVE-AD total score ≥ 8 , global rating ≥ 1 Intervention: mean BEHAVE-AD total score=16 Control: mean BEHAVE-AD total score=16
Summary	U.S.: 8/12 (67%)						Dementia diagnostic criteria: DSM IV: 9 NINCDS-ADRDA: 6 MMSE: 9 FSAT:3 Mixed: 10 Not define: 1 Severity of dementia: Moderate dementia: 5 Severe dementia: 6 Mild-severe dementia: 1	BPSD diagnostic criteria: NPI or NPI-NH: 7 BEHAVE-AD: 3 CMAI: 2 BPRS: 1 PANSS-EC: 1 CGI-S: 1 Mixed: 1

Abbreviations and scales: AD: Alzheimer's disease; BEHAVE-AD: Behavioral Pathology in Alzheimer's Disease (total score (0-75); BPRS: Brief Psychiatric Rating Scale (0-126); BPSD: Behavioral Psychological Symptoms of Dementia; CGI: Clinical Global Impression (1-7); CGI-S: Clinical Global Impression of Severity (1-7); CMAI: Cohen-Mansfield Agitation Inventory (29-203); DSM: Diagnostic and Statistical Manual of Mental Disorders; FAST: Functional Assessment Staging Test (Stages 1-7); MMSE: Mini-Mental State Examination (0-30, mild dementia=21-15, moderate dementia=11-20, severe dementia=0-10); NH: nursing home; NINCDS-ADRDA: National Institute of Neurological and Communicative Diseases and Stroke-Alzheimer's Disease and Related Disorders Association; NPI: Neuropsychiatric Inventory (0-144); NPI-NH: Neuropsychiatric Inventory-Nursing Home (2-12 for each domain. Total score for 12 domains is 24-144); PANSS-EC: Positive and Negative Syndrome Scale-Excitement Component (4-49); SIB: Severe Impairment Battery (0-100).

Table 2: Internal validity in the 12 selected trials on treating behavioral psychiatric symptoms of dementia in nursing home residents with dementia.

Publication Year, Last Name of the First Author	Power calculation	Randomization	Double blind	Concealment	Attrition (%)	Intention-to treat analysis
2008 Strein [73]	No	Yes	Yes	No	41	No
2007 Mintzer [74]	No	Yes	Yes	No	42	Yes
2007 Zhong [75]	Yes	Yes	Yes	Yes	36	Yes
2006 Mintzer [76]	Yes	Yes	Yes	Yes	35	Yes
2006 Tariot [77]	Yes	Yes	Yes	Yes	36	Yes
2005 Ballard [78]	Yes	Yes	Yes	Yes	20	Yes
2003 Brodaty [79]	Yes	Yes	Yes	No	32	Yes
2000 Street [80]	Yes	Yes	Yes	Yes	26	Yes
1999 De Deyn [81]	Yes	Yes	Yes	Yes	35	Yes
2005 Deberdt [82]	No	Yes	Yes	No	32	Yes
2004 De Deyn [83]	Yes	Yes	Yes	No	28	Yes
1999 Katz [84]	Yes	Yes	Yes	Yes	30	No
Summary: % (n/N)	75% (9/12)	100% (12/12)	100% (12/12)	58% (7/12)	33%	83% (10/12)

Table 3: Recruitment process in the 12 selected trials on treating behavioral psychiatric symptoms of dementia in nursing home residents with dementia.

Publication Year, Last Name of the First Author	Eligible fraction (%)	Enrollment fraction (%)	Recruitment fraction (%)
2008 Strein [73]	85	92	78
2007 Mintzer [74]	N/A	74	N/A
2007 Zhong [75]	81	94	77
2006 Mintzer [76]	84	100	84
2006 Tariot [77]	57	100	59
2005 Ballard [78]	N/A	N/A	N/A
2003 Brodaty [79]	80	41	33
2000 Street [80]	N/A	N/A	N/A
1999 De Deyn [81]	N/A	90	N/A
2005 Deberdt [82]	83	86	72
2004 De Deyn [83]	N/A	93	N/A
1999 Katz [84]	86	100	86

Table 4: Summary of exclusion criteria in the 12 selected RCT on treating behavioral psychiatric symptoms of dementia in nursing home residents with dementia.

Publication Year, Last Name of the First Author	Exclusion Criteria			
	Medications	Medical diseases	Psychiatric and mental illness	Other reasons for exclusion
2008 Strein [73]	Absolute contraindications for trial drug, already on antidepressants	Not mentioned	Axis I diagnosis of delirium or schizophrenia, a schizoaffective, mood, bipolar, amnesic disorder; any reversible cause of dementia,	No family or professional career informant, clinically too critical for randomization (e.g., suicide), in another trial
2007 Mintzer [74]	History of refractoriness to antipsychotics; known hypersensitivity to aripiprazole or other quinolinones; previous participation in aripiprazole trials;	Seizure disorder; unstable thyroid function; clinical significant abnormal laboratory findings;	An axis I diseases diagnosis of delirium, amnesic disorder, bipolar disorder, schizophrenia, schizoaffective disorder, mood disorder with psychotic features; Non-AD, a current major depressive episode with psychotic symptoms of hallucinations or delusions; suicidal ideation or history;	
2007 Zhong [75]	Failing to respond to a prior adequate trial of atypical antipsychotics for the treatment of agitation	Unstable medical illness (this included but was not limited to cardiovascular, renal, hepatic, hematological, endocrine, and cerebrovascular disorders); abnormal EKG that was conserved clinically significant.	History of schizophrenia, schizoaffective disorder or bipolar disorder, or agitation that was judged not related to dementia;	
2006 Mintzer [76]	Recently treated with neuroleptic injection	Medical conditions that diminish cognition; epilepsy; recent diagnosis of cancer except nonmelanoma skin cancers; unstable medical conditions; changes in prescription medications 30 days before screening; significant baseline laboratory or EKG abnormalities.	Other psychiatric disorders that produce psychotic symptoms;	Patients were withdrawn if their behavior worsened considerably, they withdraw consent, or their randomization code was broken.
2006 Tariot [77]	History of drug-induced agranulocytosis	Clinical significantly medical conditions; history of orthostasis; clinically significant EKG abnormalities;	Concurrent other Axis I DSM-IV diagnosis	
2005 Ballard [78]	Known to be sensitive to cholinesterase inhibitor or antipsychotics	Advanced, severe, progressive, or unstable disease that might interfere with efficacy or put the patient at special risk, severe, unstable, or poorly controlled medical conditions, bradycardia (<50), sick sinus syndrome, or conduction defect, current diagnosis of active uncontrolled peptic ulceration within the past three months; and clinically significant urinary obstruction		Disability that might prevent patients from completing study procedures

2003 Brodaty [79]	Administration of a depot neuroleptic within 2 treatment cycles, a history of neuroleptic malignant syndrome or an allergic reaction to neuroleptic drugs, history of failure to respond to risperidone treatment of at least 4 weeks' duration and participation in clinical trial(s) with any investigational drugs during the 4 weeks preceding selection	Medical or neurologic conditions other than dementia that diminish cognitive function; clinically uncontrolled organic disease, clinically relevant laboratory abnormalities,	Other type dementia (not AD or vascular dementia), major depression within the last 6 months, other psychiatric disorders that could have accounted for observed psychotic disturbances, a history of tardive dyskinesia	
2000 Street [80]			A history of a DSM-IV Axis I disorder (e.g., schizophrenia, recurrent depression), any neurological condition other than AD that could contribute to psychosis or dementia, a Mini-Mental State Examination score of greater than 24	Bedridden status
1999 De Deyn [81]	Administration of a depot neuroleptic within one treatment cycle of Visit 1; history of allergic reaction to neuroleptics or history of neuroleptic malignant syndrome; or participation in clinical trial(s) with investigational drugs during the 4 weeks preceding this trial		Other conditions that diminish cognitive function; other psychiatric disorders; clinical relevant organic or neurologic disease; EKG or laboratory abnormalities	
2005 Deberdt [82]	Atypical antipsychotic use was disallowed within 30 days, lithium or anticonvulsant use within 2 weeks before the placebo/ washout period. Oral conventional antipsychotic use was allowed up to 3 days before randomization.	Parkinson disease, Pick disease,	Low-body dementia, front temporal dementia, Mini-mental State Exam score <5 or >24,	
2004 De Deyn [83]	Cholinesterase inhibitory therapy was allowed to continue into the study if the patients had been on a stable dose prior to Visit 1.		A diagnosis of current primary mood disorder or other Axis I disorder with onset prior to diagnosis of AD	
1999 Katz [84]		Medical or neurologic conditions that diminish cognitive function, dementia related to infection with the human immunodeficiency virus or substance-induced persistent dementia	Untreated reversible dementia, delirium or amnesic disorder, a psychiatric diagnosis that could have accounted for the observed psychotic disturbances	

Abbreviations: AD: Alzheimer's disease; DSM: Diagnostic and Statistical Manual of Mental Disorder

measure for efficacy of atypical antipsychotics or placebo across the 12 trials. We reported qualitative descriptions and estimates of the intervention effect among the 12 RCTs and followed one previous systematic review [41] to examine the consistency of the intervention effects by point estimates across the 12 trials. Because we did not perform meta-analysis, we did not have I^2 and CIs.

Planned methods of analysis

We decided not to perform a meta-analysis because the diagnostic criteria of dementia and BPSD, BPSD outcome measurements, and rating scales varied across the 12 trials. Instead, we provided qualitative descriptions and estimates following one previous systematic review without a meta-analysis [41], which is the only systematic review focusing on NH patients with BPSD. Descriptive analyses were performed in SPSS (Version 18). Data was analyzed by one author (HYC) and discussed with another author (TNH).

Risk of bias across studies

Because we decided not to perform a meta-analysis, risk of bias across the 12 RCTs was not assessed. Missing studies, missing outcomes, and detection of missing information were not examined.

We used descriptive approach to describe the differences of risk of bias across the 12 trials which are summarized in Table 2.

Analysis on placebo effects

The placebo effect is a common phenomenon in pharmacological trials of treating depression and other conditions [66-72] and was assessed across the 12 RCTs in this study.

Others

Diagnostic criteria, outcome measurements, and scales of dementia and BPSD were obtained via searching the original studies and a standard textbook [61] and are listed in the footnote of Tables 1 and 5. Secondary outcomes were also collected and are summarized in Table 5, which is not the focus of our systematic review.

Results

Study selection

From PubMed we found 1469 relevant citations using the pre-defined search criteria. Of those, 1449 were excluded because they did not meet the inclusion criteria. Twenty original papers were fully

Table 5: Interventions and outcomes in the 12 selected trials on treating behavioral psychiatric symptoms of dementia in nursing home residents with dementia.

Publication Year, Last Name of the First Author	Primary Intervention	Primary Outcome (Intervention vs Control)	Adverse effects	List of Secondary Outcome Measurements
2008 Strein [73]	-1 week washout period -Aripiprazole 2 mg/day up to 15 mg/day based on clinician judgment -Try duration: 10 weeks	-No statistical difference of NPI-NH psychosis subscale (6 vs 6 on a scale of 2-12) and CGI-S (4 vs 4 on a scale of 1-7) -Intervention effect: decrease NPI-NH by 43% and CGI-S by 13% -Placebo effect: decrease NPI-NH by 43% and CGI-S by 10%	Multiple adverse effects (accidental injuries, somnolence, ulcer skin, urinary tract infection, vomiting, stroke, EPS, death) in both interventional and control groups	NPI-NH total, BPRS total, psychosis, and core, CMAI, CSDD, NPI-NH total and psychosis of caregivers distress, ADCS-ADL-SEV, MMSE,
2007 Mintzer [74]	-Aripiprazole at 3 fixed dose: 2 mg/day, 5 mg/day and 10 mg/day -Trial duration: 10 weeks	Statistical difference of NPI-NH psychosis subscale (8.6 vs 10 on a scale of 2-12). It might not have clinical significance. Interventional effect in 10 mg group: reduce NPI-NH subscale by 44% Placebo effect: reduce NPI-NH subscale by 34%	Multiple adverse effects (abnormal gait, weight loss, confusion, vomiting, insomnia, anorexia, constipation etc.) in both intervention and placebo groups. However, 10 mg Aripiprazole group had 15% serious adverse effects and placebo group had 8%.	NPI-NH total, CGI-S, BPRS core, subscale and total, CMAI, MMSE, CGI-I
2007 Zhong [75]	-Quetiapine: started at 25 mg/day and titrated to 100 mg or 200 mg/day -Patients unable to tolerate titration were excluded -Trial duration: 10 weeks	No statistical difference of PANSS-EC (17.3 vs 18.9 on a scale of 4-49) Intervention effect: reduce PANSS-EC by 26% in 200 mg quetiapine group. Placebo effect: reduce PANSS-EC by 17%.	No difference between intervention and placebo groups. Multiple adverse effects (fall, vomiting, EPS, gait abnormalities, weight decreased etc.) in both intervention and placebo groups.	Response rate to PANSS-EC, NPI-NH total and 4 subscales, CMAI total and 3 subscales, MMSE, CGI-C, ADCS
2006 Mintzer [76]	-Run-in phase for 1 week -Risperidone: started at 0.5 mg and can be titrated to 1.5 mg/day -Trial duration: 8 weeks	No statistical difference of BEHAVE-AD and CGI-C. Intervention and placebo effect: unable to calculate (the results were presented in Figure).	Multiple adverse effects (fall, insomnia, EPS, somnolence etc.) in both intervention and placebo groups. Intervention group had 14% serious adverse effects and placebo group had 13%.	BEHAVE-AD total and 5 subscales and a global item, CGI response
2006 Tariot [77]	-Wash-out ≥ 48 hours -Quetiapine: started at 25 mg/day and titrated to 100 mg/day -Both Haldol and placebo as control -Trial duration: 10 weeks	No statistical difference of BPRS total (31 vs 32 on a scale of 0-75) and CGI-S (4 vs 4 on a scale of 1-7) Intervention effect: reduce BPRS total by 23% and CGI-S by 13%, Placebo effect: reduce BPRS total by 17% and CGI-S by 10%.	Multiple adverse effects (falls, somnolence, vomiting etc.) in both intervention and placebo groups.	3 BPRS subscales, NPI-NH2 and agitation, MMSE, MOSES
2005 Ballard [78]	-Quetiapine: (dose??) -Rivastigmine and placebo as control -Trial duration: 6 weeks	No statistical difference of CMAI agitation (55.1 vs 50.2 on a scale of 29-203) Intervention group: reduce CMAI-Agitation by 7% Placebo effect: reduce CMAI-Agitation by 11%	Not fully reported	SIB
2003 Brodaty [79]	-Risperidone: 0.5 mg/day up to 2mg/day -Trial duration: 12 weeks	Statistical significance of CMAI total aggression (decline of 8.5 vs 4.5). It might not be clinically significant on a scale of 14-98. Intervention effect: reduce CMAI-total agitation score by 25% Placebo effect: reduce CMAI-total agitation score by 14%	Multiple adverse effects (fall, somnolence, vomiting etc.) in both intervention and placebo groups. Intervention group had 3.4% serious adverse effects and placebo had 2.4%.	CMAI total aggression and 2 subscales, CMAI total non-aggression and 2 subscales, BEHAVE-AD total and 8 subscales, CGI-C, CGI-S, MMSE, FAST
2000 Street [80]	-Washout and placebo lead-in for 4-14 days -Olanzapine: 5 mg/day, 10 mg/day, 15 mg/day. Patients unable to tolerate the assigned treatment were excluded -Trial duration: 6 weeks	Statistically significance of NPI-NH core total between 5 or 10 mg/day and placebo (declined by 7.6, 6.1 vs 3.7 on a scale of 0-36). No difference between 15 mg/day and placebo. Intervention effect: reduce NPI-NH core total by 35% in 15 mg Olanzapine group Placebo effect: reduce NPI-NH core total by 15%.	Multiple adverse effects (accidental injuries, somnolence, pain, abnormal gait etc.) in both intervention and placebo effect. No difference of EPS between intervention and placebo groups.	NPI-NH total, NPI-NH psychosis total and 5 subscales, BPRS total and 2 subscales, MMSE
1999 De Deyn [81]	-Washout for 1 week -Risperidone: started at 0.25 mg/day and maximum dose up to 4 mg/day -Haloperidol and placebo as control -Trial duration: 12 weeks	Difference of BEHAVE-AD total score did not reach statistical significance at the end of 12 weeks based on at least 30% reduction of BEHAVE-AD total score (72% vs 61%, p=0.13). Intervention effect: 72% clinical improvement on BEHAVE-AD Placebo effect: 61% clinical improvement on BEHAVE-AD	Intervention group had 76% adverse effects and placebo group had 72.8%. No difference of serious or severe adverse effects between intervention and placebo groups.	BEHAVE-AD aggressiveness, BEHAVE-AD global rating, CMAI 3 subscales, CGI-S

2005 Deberdt [82]	-Washout/placebo period from 3-12 days -Olanzapine: flexible dose from 2.5 mg/day to maximum dose of 10 mg/day -Risperidone and placebo as control -Trial duration: 10 weeks	No statistical difference of NPI or NPI-NH psychosis total (7 vs 6.4 and CGI-S psychosis scale (3.3 vs 3.3) Intervention effect: reduce NPI psychosis total by 35% and CGI-Severity of Psychosis by 18% Placebo effect: reduce NPI Psychosis Total by 42% and CGI-Severity of Psychosis by 20%	Multiple adverse effects (somnia, agitation, accidental injury etc.) in both intervention and placebo groups.	NPI caregiver total, NPI total, BPRS total, CMAI aggression, PDS, CSDD
2004 De Deyn [83]	-A placebo lead-in phase up to maximum 14 days -Olanzapine: 1 mg, 2.5 mg, 5 mg, 7.5 mg/day. 5 and 7.5 mg/day was as intervention -Started at 1 mg/day. Patients unable to tolerate the assigned olanzapine or placebo were excluded. -Trial duration: 10 weeks	No statistically difference of NPI or NPI-NH psychosis total and CGI-C Intervention effect: reduce NPI-NH psychosis total by 64% in 7.5 mg Olanzapine group Placebo effect: reduce NPI-NH psychosis total by 52%	No differences of adverse effects between intervention and placebo	NPI/NH total with 8 subscales, BPRS total with 2 subscales, CGI-S, MMSE, SIB
1999 Katz [84]	-Washout for 3-7 days -Risperidone: 0.5 mg, 1 mg and 2 mg/day -Trial duration: 12 weeks	Statistically difference of BEHAVE-AD total score between 1 mg/day or 2 mg/day and placebo (decline by 7 and 8 vs 5 on a scale of 0-75). It is considered to be less clinically significant. Intervention effect: reduce BEHAVE-AD total score by 38-67% (mean=50%) Placebo effect: reduce BEHAVE-AD total score by 31%	Multiple adverse effects (injury, somnolence, fall etc.) in both intervention and placebo groups.	BEHAVE-AD 2 subscales, CMAI verbal and physical aggression, CGI, FAST, MMSE
Summary	1). Atypical agents tested: Aripiprazole (n=2) Olanzapine (n=3) Quetiapine (n=3) Risperidone (n=5) 2). 67% (8/12) trials had washout or run-in period. 3). Dosing: 92% (11/12) trials used titration.	33 % (4/12) trials statically significantly reduced BPSD	1). 92% (11/12) trials reported adverse effects. 2). Intervention groups tended to have similar adverse effects to placebo groups. 17% (2/12) trials showed more serious adverse effects in intervention groups than placebo groups.	1). All trials reported at least one or more secondary outcomes

Abbreviations and scales: ADCS-ADL: Alzheimer’s Disease Cooperative Study – Activities of Daily Living Scale (0-78); BEHAVE-AD: Behavioral Pathology in Alzheimer’s Disease (total score (0-75); BPRS: Brief Psychiatric Rating Scale (0-126); BPSD: Behavioral Psychological Symptoms of Dementia; CGI: Clinical Global Impression (1-7); CGI-C: Clinical Global Impression of Change; CGI-I: Clinical Global Impression-Global Improvement (1-7); CGI-S: Clinical Global Impression of Severity (1-7); CMAI: Cohen-Mansfield Agitation Inventory (29-203); CSDD or CS: Cornell scale for depression in dementia (0-38); EPS: Extrapyramidal symptoms; FAST: Functional Assessment Staging Test (Stages 1-7); MMSE: Mini-Mental State Examination (0-30, mild dementia=21-15, moderate dementia=11-20, severe dementia=0-10); MOSES: Multidimensional Observation Scale for Elderly Subjects-Social Activities Subscale; NH: nursing home; NPI-NH: Neuropsychiatric Inventory-Nursing Home (2-12 for each domain. Total score for 12 domains is 24-144); PANSS-EC: Positive and Negative Syndrome Scale-Excitement Component (4-49); PDS: Progressive Deterioration Scale (0-100); SIB: Severe Impairment Battery (0-100).

reviewed. Twelve papers [73-84] met the inclusion criteria (Figure 1). Searches through Cochrane review, the National Guideline Clearinghouse database, previously published systematic reviews [31-43], and a hand search for references did not find additional papers that met the inclusion criteria. We identified more RCTs on our PICOS question than did previous systematic reviews [31-43]. The excluded RCTs and the reasons for exclusion from our systematic review are not reported here, but are available as a supplemental file online.

Study characteristics

The 12 selected RCTs are summarized in Table 1. All 12 RCTs [73-84] were published between 1999 and 2008. Nine of the 12 RCTs were conducted only in the NH setting [73-81]. Three of the 12 RCTs were conducted either in NH and in clinics [82], or in NH and in hospitals [83,84]. Most trials were sponsored by industry and conducted in the United States. All trials except one [76] used either the Diagnostic and Statistical Manual of Mental Disorders (DSM) [73-75,77,79-84] or the National Institute of Neurological and Communicative Diseases and Stroke-Alzheimer’s Disease and Related Disorders Association (NINCDS-ADRDA) criteria [78,80], or both [75], to diagnose

dementia. The Mini-Mental State Examination (MMSE) was used in 11 trials [73-77,79-84] to assess the severity of dementia and the Functional Assessment Staging Test (FAST) was used in one trial [78]. One trial used both MMSE and FAST [84]. Severity of dementia ranged from mild to severe across the 12 trials. The instruments used to measure BPSD as the primary outcome varied across the 12 trials (Table 1) and included the Neuropsychiatric Inventory-Nursing Home (NPI-NH) [73,74,80,83], the Behavioral Pathology in Alzheimer’s Disease (BEHAVE-AD) [76,81,84], the Cohen-Mansfield Agitation Inventory (CMAI) [78,79], and the Positive and Negative Syndrome Scale-Excitement Component (PANSS-EC) [75]. The Neuropsychiatric Inventory (NPI) and the Brief Psychiatric Rating Scale (BPRS) were used together with other instruments [77,81]. BPSD severity was not defined due to lacking well-accepted scale criteria. However, the scores of BPSD severity are generally at low levels of the BPSD scales across the 12 trials. Among the total 4352 enrolled subjects across the 12 trials, the age of the enrolled participants ranged from 77 to 84 years old. Fifty eight percent to 72% of the participants were female and 80% to 100% of participants were white. The sample size of the enrolled participants ranged from 62 to 652. Among the enrolled participants, 2924 received an atypical

antipsychotic agent as the intervention and 1425 received placebo (three research subjects dropped out before the interventions from one trial) [83]. Sixty seven percent and 70% completed the trial in the intervention and placebo groups, respectively. Five of the 12 trials [74,79,81-83] did not report the number of individuals screened. Diagnostic instruments of dementia and BPSD measurements and scales are listed in the footnote of Tables 1 and 5. Finally, IRB approval and informed consent were obtained among all 12 trials.

Risk of bias in individual studies

The internal validity of the 12 selected RCTs is summarized in Table 2. All trials used randomization and double blinding (performance bias and detection bias). Seventy five percent of the trials included power calculations. Fifty eight percent of trials reported concealment (selection bias). The attrition rate ranged from 20% to 42% across the 12 trials (attrition bias). Power calculation was performed in 75% of the 12 trials. Intention-to-treat analysis was used in 83% of the 12 trials.

Applicability in individual studies

The recruitment process indicators are summarized in Table 3. Seventy nine percent of the trials reported eligible fractions, which ranged from 57% to 86%. Eighty seven percent of trials reported enrollment fractions, which ranged from 41% to 100%. Sixty nine percent of trials reported recruitment fractions. Among reported trials, recruitment fractions ranged from 33% to 86%. Exclusion criteria included certain medications, co-existing medical and other psychological diagnoses, and other reasons and are summarized in Table 4.

Heterogeneity among the selected trials

Heterogeneity includes three type of diversity [61]. 1) Clinical diversity (clinical heterogeneity): Table 1 shows all research participants had dementia and BPSD. However, different atypical antipsychotics were used. Severity of dementia and BPSD and BPSD outcome measurements varied across the 12 trials. 2) Methodological diversity: All trials used randomized controlled trial design. However, there were significant variations of risk of bias including concealment and attrition rate across the 12 trials (Table 2). 3) Statistical heterogeneity: The intervention effects from atypical antipsychotics varied across the 12 trials (Table 5). We did not perform a meta-analysis, and so summary of statistical heterogeneity are not available.

Synthesis of results

We did not perform a meta-analysis for the intervention effects. Therefore, we reported qualitative descriptions and estimates of intervention effect among the 12 RCTs following one previous systematic review [41] to examine the consistency of the intervention effects among the 12 trials.

The atypical antipsychotics tested among the 12 RCTs included risperidone (N=4) [76,79,81,84], quetiapine (N=3) [75,77,78], olanzapine (N=3) [80,82,83], and aripiprazole (N=2) [73,74]. The four RCTs that demonstrated a statistically significant BPSD reduction utilized aripiprazole (N=1) [74], risperidone (N=2) [79,84], and olanzapine (N=1) [80]. The three trials with quetiapine [75,77,78] and five other trials [73,76,81-83] did not demonstrate statistically significant BPSD reduction. There was inconsistent BPSD

reduction among trials testing risperidone, quetiapine, olanzapine, and aripiprazole. The intervention effects on BPSD reduction varied from 7% to 72% and were inconsistent among the 12 trials. There were no statistically significant differences in the number of enrolled participants or in attrition rates between the eight negative and four positive trials (unadjusted).

All trials reported adverse effects, which were similar between the intervention and placebo groups. A brief list of adverse profiles is presented in Table 5, which is not the focus of this systematic review. Most trials tended to report multiple secondary outcomes in Table 5, which is not the focus of this systematic review.

Risk of bias across studies

Because we decided not to perform a meta-analysis, we did not analyze the risk of bias across the 12 trials. Instead, we reported the risk of bias for individual studies above (Table 2). Missing studies, missing outcomes, and detection of missing information were not examined in this systematic review.

Placebo effects

The effects of placebo on BPSD reduction were reported in the 11 trials except one [74] for which we were unable to calculate a placebo effect. In this trial, the author used a figure for the primary outcome [74]. The placebo effect varied from 14% to 34% among the four RCTs that showed statistically significant BPSD reduction [74,79,80,84] and from 10% to 61% among the seven RCTs that showed no statistically significant BPSD reduction [73,75,77,78,81-83] (Table 5).

Discussion

Though both the FDA and CMS have released warnings and recommendations regarding cautious use of atypical antipsychotics due to the risk of adverse effects [19,20], these agents are still widely used in NH residents [11-17]. Based on the limited evidence of efficacy of atypical antipsychotics, this systematic review further supports these recommendations. In this review, we addressed the following four questions.

First, can atypical antipsychotics reduce BPSD among NH residents? Our major findings indicate that atypical antipsychotics did not reduce BPSD among NH residents in eight of the 12 RCTs. This is consistent with a previous systematic review that focused on NH residents [41]. However, it is inconsistent with previous systematic reviews that were not specific to NH settings [31-40]. This suggests that the results of systematic reviews from the non-NH setting might not be relevant to the NH population. One of reasons for the negative results in these eight RCTs may be the low level of BPSD scales in these participants, despite the severity of BPSD not being defined. We have provided the BPSD measurements and rating scales on the footnote of Tables 1 and 5 which may help the readers to assess the BPSD severity. For example, the mean score of BPSD in the intervention arm in one trial was 40 on a scale of 0-126 [77], while in another study, the mean score of the CMAI in the intervention arm was 59 on a scale of 29-203 [78]. These trials may have shown greater efficacy if investigators had recruited participants with higher levels of BPSD. However, four positive trials showed a statistically significant BPSD reduction. These inconsistent findings reduce our confidence in the intervention of atypical antipsychotics for BPSD

reduction. Whether some patients might benefit from the use of atypical antipsychotic agents needs to be confirmed in NH residents with BPSD. In this systematic review, there was no difference in adverse effects between the intervention and the placebo groups, which is most likely due to the small sample size. NH residents are often frail and have multiple co-existing conditions [85] and tend to take multiple medications [86-88], while older adults in general are sensitive to antipsychotic agents [89]. Due to the limited evidence for the use of atypical antipsychotics, the potential for adverse effects [1,18-30], and risks of polypharmacy [86-88], we fully understand the concerns, warnings, and regulations from the FDA and CMS [18-20]. Atypical antipsychotics should be used cautiously for NH residents with BPSD.

Second, are the RCTs examining atypical antipsychotics to treat NH residents with BPSD valid, i.e. is the internal validity good? The internal validity is defined as whether the study results in true findings and minimizes systematic error [61-64]. The internal validity is critically important in RCTs [61-64]. Randomization and blinding among the 12 RCTs were well reported. However, we are very concerned about the high attrition rates among the 12 trials, which may result in attrition bias [61]. High attrition rates in RCTs among the elderly population are expected because older participants might die or drop out due to pre-existing conditions or side effects, but researchers have proposed multiple ways to retain elderly participants [90]. We are also concerned with the low allocation concealment among the 12 trials, which can be an indication of selection bias [61]. A good trial of antipsychotics to treat NH residents with BPSD with a high retention of participants and high allocation concealment should be conducted to reduce the bias and to improve the internal validity.

Third, what is the role of the placebo effect of atypical antipsychotics in the treatment of BPSD in NH residents? Placebo effects in the treatment of depression and other conditions are well documented in RCTs [61,66-72], but the mechanism of the placebo effect is complex and unknown [68-70]. To our knowledge, we may be the first to examine the placebo effect in both the negative and positive trials examining treatment of BPSD in NH residents. The positive trials had a larger placebo effect than did the negative trials, which indicates that the intervention effect in the positive trials was not less likely due to a small placebo effect. While previous studies that showed significant placebo effects often used patients' self-reported subjective symptoms and research subjects' responses to interviews, such as with pain or depression [61,66,71,72], BPSD across the 12 trials was observed and measured by researchers. Therefore, placebo effects in these trials cannot arise from the NH residents themselves, but more likely arise from the researchers. Another potential mechanism of the placebo effect could be due to the therapeutic effect of the interaction between the researchers and participants, which needs to be tested. Finally, these trials lasted from six to 12 weeks. BPSD natural history is toward a reduction over time, which may also explain the observed placebo effect.

Fourth, are the results of the RCTs applicable to the population in the real world, i.e. is the external validity good [91-93]? External validity has been an issue in conducting RCTs. For example, research subjects from RCTs in treating depression may have only represented the minority of the patients in the real world [94-98]. Our results showed that the research subjects across the 12 RCTs in treating BPSD

in NH patients with dementia were not representative of the real-world patients because many participants were excluded from the trials, including those on certain medications, or having co-existing medical diseases and other psychological conditions. The highly selected populations in these RCTs are likely significantly different from real NH patients with dementia who often have multiple co-existing conditions [99-101]. In addition, the participant recruitment process across the 12 RCTs was under-reported, and therefore it is difficult for the readers to know from where the enrolled research subjects came. Taken together, all these factors could significantly reduce the external validity of these 12 RCTs. Better reporting methods and better recruitment processes are needed in RCTs examining treatment of NH residents with BPSD. Recruiting typical NH patients via a pragmatic RCT design should be used to improve the applicability and guide daily practice [102-105].

Heterogeneity is a crucial part of the assessment of RCTs and other types of studies [61]. Heterogeneity includes three types of diversity [61]. 1) Clinical diversity (clinical heterogeneity) is present among the 12 trials because four different atypical antipsychotics (risperidone, quetiapine, olanzapine, and aripiprazole) with different doses were used, and because the severity of dementia and BPSD and BPSD outcome measurements varied across the 12 trials. 2) Methodological diversity is present among the 12 trials because of significant variations of the risk of bias including concealment and attrition rates. 3) Statistical heterogeneity is present among the 12 trials because the intervention effects from four atypical antipsychotics varied. We did not perform a meta-analysis, so a summary of statistical heterogeneity are not available. We are concerned with the heterogeneity of the 12 studies analyzed, as it reduces the internal and external validity of the findings.

The discussion of other findings is worthwhile. Based on our own experience, primary care providers are often uncomfortable using research instruments such as BPSD outcome measurements as they are less commonly used in their daily primary practice. This could limit their understanding and consequent application of the results from these RCTs. Additionally, the details of BPSD outcome measurements and the rating scales were not fully reported in previous systematic reviews. This systematic review provides the various BPSD outcome measurements and scales in order to help primary care providers understand these measurements and results of the 12 RCTs.

Since the Cochrane Handbook for Systematic Reviews of Intervention was published in 2008 [61], The Grades of Recommendations, Assessment, Development, and Evaluation (GRADE) has been rapidly evolved to help systematic reviewers, clinical practice guideline developers, and health technology assessments [106]. The first part of GRADE (evidence profiles) is well developed to rate the quality of evidence-based risk of bias, publication bias, imprecision (random error), inconsistency, and indirectness and to classify the confidence of effect estimates into four categories from high, moderate, low to very low quality of evidence [106-115]. The Evidence-based Practice Center (EPC) program funded by the U.S. Agency for Health Research and Quality (AHRQ) has updated their grading system accordingly based on GRADE [116,117]. Because we were not familiar with the evolution of GRADE since 2010, we did not integrate GRADE into our systematic review. However, we would

like to briefly summarize the quality of evidence for the 12 selected RCTs in our systematic review based on the latest GRADE framework as following. 1) Risk of bias. It is good to have blindness and report primary and secondary outcomes among the 12 trials. However, the low concealment rate (58%) and high drop-out (20-42%) among the 12 trials indicates significant risk of bias. 2) Publication bias. The small number of selected RCTs, the small sample size of some RCTs, and possible missed studies including unknown unpublished studies increases the risk of publication bias in our systematic review. 3) Imprecision. CIs and optimal information size (OIS) were not obtained, which indicates the imprecision might be present among the 12 trials. 4) Inconsistency. The point estimates of intervention effects on BPSD reduction varied widely from 7% to 72% among the 12 trials, which is considered an inconsistency. We did not perform a meta-analysis and unable to provide information on statistical heterogeneity and I^2 . The inconsistency in our systematic review is uncertain. 5) Indirectness. Highly selected research participants (poor applicability), different atypical antipsychotics with different doses and trial durations, and participants from academic settings among the 12 trials indicate indirectness. Taken together, we rate the quality of evidence from the 12 trials as low.

We admit our systematic review has several limitations. GRADE framework was found to be reproducible [118] and recently used to assess the quality of evidence via the format of evidence profile and summary of findings [118,119]. We did not integrate this format with our systematic review in the beginning. In our opinion, this format will be fully accepted by the Cochrane Handbook for systematic review of interventions in the future. We did not write a full search protocol, however we did have a search plan. A true pre-specification of methods was not done because one of the authors wrote an invited chapter on BPSD and knew few RCTs on BPSD. This review had a small number of RCTs. A meta-analysis was not performed due to the different diagnostic criteria and severity scales of dementia and BPSD, as well as different BPSD outcome measurements and rating scales among the 12 trials. Non-English papers were excluded, and unpublished data were not collected. PubMed was used to search and identify relevant papers, but other datasets were not used. Despite our extensive search terms and search of Cochrane reviews, the National Guideline Clearinghouse database, previously published systematic reviews, and hand search, we may have missed some relevant RCTs which results in publication bias. Missing studies, missing outcomes, and detection of missing information were not fully examined.

Conclusions

There is limited and inconsistent evidence to demonstrate the efficacy of atypical antipsychotics in the treatment of behavioral and psychological symptoms of dementia in nursing home residents. In addition, there are concerns about the risk of bias, applicability, clinical heterogeneity, and methodological diversity among the 12 trials. We are less confident in the intervention effects of atypical antipsychotics on BPSD in nursing home patients and therefore do not recommend routinely prescribing atypical antipsychotics for this indication.

References

1. <http://www.ipa-online.net/pdfs/1BPSDFinal.pdf>

- Zuidema S, Koopmans R, Verhey F (2007) Prevalence and predictors of neuropsychiatric symptoms in cognitively impaired nursing home patients. *J Geriatr Psychiatry Neurol* 20: 41-49.
- Wetzels R, Zuidema S, Jansen I, Verhey F, Koopmans R (2010) Course of neuropsychiatric symptoms in residents with dementia in long-term care institutions: a systematic review. *Int Psychogeriatr* 22: 1040-1053.
- Selbæk G, Engedal K, Bergh S (2013) The prevalence and course of neuropsychiatric symptoms in nursing home patients with dementia: a systematic review. *J Am Med Dir Assoc* 14: 161-169.
- Pelletier IC, Landreville P (2007) Discomfort and agitation in older adults with dementia. *BMC Geriatr* 7: 27.
- Zwijzen SA, Kabboord A, Eefsting JA, Hertogh CM, Pot AM, et al. (2014) Nurses in distress? An explorative study into the relation between distress and individual neuropsychiatric symptoms of people with dementia in nursing homes. *Int J Geriatr Psychiatry* 29: 384-391.
- Kurasawa S, Yoshimasu K, Washio M, Fukumoto J, Takemura S, et al. (2012) Factors influencing caregivers' burden among family caregivers and institutionalization of in-home elderly people cared for by family caregivers. *Environ Health Prev Med* 17: 474-483.
- de Rooij AH, Luijkx KG, Declercq AG, Emmerink PM, Schols JM (2012) Professional caregivers' mental health problems and burnout in small-scale and traditional long term care settings for elderly people with dementia in the Netherlands and Belgium. *J Am Med Dir Assoc* 13: e7-e11.
- Gaugler JE, Wall MM, Kane RL, Menk JS, Sarsour K, et al. (2011) Does caregiver burden mediate the effects of behavioral disturbances on nursing home admission? *Am J Geriatr Psychiatry* 19: 497-506.
- Briesacher BA, Tjia J, Field T, Peterson D, Gurwitz JH (2013) Antipsychotic use among nursing home residents. *JAMA* 309: 440-442.
- Richter T, Mann E, Meyer G, Haastert B, Köpke S (2012) Prevalence of psychotropic medication use among German and Austrian nursing home residents: a comparison of 3 cohorts. *J Am Med Dir Assoc* 13: e7-e187.
- Zuidema SU, de Jonghe JF, Verhey FR, Koopmans RT (2011) Psychotropic drug prescription in nursing home patients with dementia: influence of environmental correlates and staff distress on physicians' prescription behavior. *Int Psychogeriatr* 23: 1632-1639.
- Chen Y, Briesacher BA, Field TS, Tjia J, Lau DT, et al. (2010) Unexplained variation across US nursing homes in antipsychotic prescribing rates. *Arch Intern Med* 170: 89-95.
- Kamble P, Chen H, Sherer JT, Aparasu RR (2009) Use of antipsychotics among elderly nursing home residents with dementia in the US: an analysis of National Survey Data. *Drugs Aging* 26: 483-492.
- Lövheim H, Sandman PO, Karlsson S, Gustafson Y (2009) Changes between 1982 and 2000 in the prevalence of behavioral symptoms and psychotropic drug treatment among old people with cognitive impairment in geriatric care. *Int Psychogeriatr* 21: 941-948.
- Kamble P, Chen H, Sherer J, Aparasu RR (2008) Antipsychotic drug use among elderly nursing home residents in the United States. *Am J Geriatr Pharmacother* 6: 187-197.
- Lövheim H, Sandman PO, Kallin K, Karlsson S, Gustafson Y (2006) Relationship between antipsychotic drug use and behavioral and psychological symptoms of dementia in old people with cognitive impairment living in geriatric care. *Int Psychogeriatr* 18: 713-726.
- <http://oig.hhs.gov/oei/reports/oei-07-08-00150.pdf>
- <http://www.fda.gov/safety/medwatch/safetyinformation/safetyalertsforhumanmedicalproducts/ucm150688.htm>.
- <https://www.cms.gov/Regulations-and-Guidance/Guidance/Transmittals/downloads/R22SOMA.pdf>
- Morley JE (2012) Antipsychotics and dementia: a time for restraint? *J Am Med Dir Assoc* 13: 761-763.
- Kales HC, Kim HM, Zivin K, Valenstein M, Seyfried LS, et al. (2012) Risk

- of mortality among individual antipsychotics in patients with dementia. *Am J Psychiatry* 169: 71-79.
23. Pariente A, Fourrier-Réglat A, Ducruet T, Farrington P, Béland SG, et al. (2012) Antipsychotic use and myocardial infarction in older patients with treated dementia. *Arch Intern Med* 172: 648-653.
24. Hill KD, Wee R (2012) Psychotropic drug-induced falls in older people: a review of interventions aimed at reducing the problem. *Drugs Aging* 29: 15-30.
25. Rochon PA, Normand SL, Gomes T, Gill SS, Anderson GM, et al. (2008) Antipsychotic therapy and short-term serious events in older adults with dementia. *Arch Intern Med* 168: 1090-1096.
26. Gill SS, Bronskill SE, Normand SL, Anderson GM, Sykora K, et al. (2007) Antipsychotic drug use and mortality in older adults with dementia. *Ann Intern Med* 146: 775-786.
27. Bozat-Emre S, Doupe M, Kozyrskyj AL, Grymonpre R, Mahmud SM (2014) Atypical antipsychotic drug use and falls among nursing home residents in Winnipeg, Canada. *Int J Geriatr Psychiatry*.
28. Gill SS, Rochon PA, Herrmann N, Lee PE, Sykora K, et al. (2005) Atypical antipsychotic drugs and risk of ischaemic stroke: population based retrospective cohort study. *BMJ* 330: 445.
29. Wang PS, Schneeweiss S, Avorn J, Fischer MA, Mogun H, et al. (2005) Risk of death in elderly users of conventional vs. atypical antipsychotic medications. *N Engl J Med* 353: 2335-2341.
30. Schneider LS, Dagerman KS, Insel P (2005) Risk of death with atypical antipsychotic drug treatment for dementia: meta-analysis of randomized placebo-controlled trials. *JAMA* 294: 1934-1943.
31. Ballard C, Waite J (2006) The effectiveness of atypical antipsychotics for the treatment of aggression and psychosis in Alzheimer's disease. *Cochrane Database Syst Rev* 25: CD003476.
32. Maher AR, Maglione M, Bagley S, Suttrop M, Hu JH, et al. (2011) Efficacy and comparative effectiveness of atypical antipsychotic medications for off-label uses in adults: a systematic review and meta-analysis. *JAMA* 306: 1359-1369.
33. Herrmann N, Lanctôt KL (2007) Pharmacologic management of neuropsychiatric symptoms of Alzheimer disease. *Can J Psychiatry* 52: 630-646.
34. Schneider LS, Dagerman K, Insel PS (2006) Efficacy and adverse effects of atypical antipsychotics for dementia: meta-analysis of randomized, placebo-controlled trials. *Am J Geriatr Psychiatry* 14: 191-210.
35. Carson S, McDonagh MS, Peterson K (2006) A systematic review of the efficacy and safety of atypical antipsychotics in patients with psychological and behavioral symptoms of dementia. *J Am Geriatr Soc* 54: 354-361.
36. De Deyn PP, Katz IR, Brodaty H, Lyons B, Greenspan A, et al. (2005) Management of agitation, aggression, and psychosis associated with dementia: a pooled analysis including three randomized, placebo-controlled double-blind trials in nursing home residents treated with risperidone. *Clin Neurol Neurosurg* 107: 497-508.
37. Sink KM, Holden KF, Yaffe K (2005) Pharmacological treatment of neuropsychiatric symptoms of dementia: a review of the evidence. *JAMA* 293: 596-608.
38. Tariot PN, Proffenno LA, Ismail MS (2004) Efficacy of atypical antipsychotics in elderly patients with dementia. *J Clin Psychiatry* 65 Suppl 11: 11-15.
39. Snowden M, Sato K, Roy-Byrne P (2003) Assessment and treatment of nursing home residents with depression or behavioral symptoms associated with dementia: a review of the literature. *J Am Geriatr Soc* 51: 1305-1317.
40. Kindermann SS, Dolder CR, Bailey A, Katz IR, Jeste DV (2002) Pharmacological treatment of psychosis and agitation in elderly patients with dementia: four decades of experience. *Drugs Aging* 19: 257-276.
41. Seitz DP, Gill SS, Herrmann N, Brisbin S, Rapoport MJ, et al. (2013) Pharmacological treatments for neuropsychiatric symptoms of dementia in long-term care: a systematic review. *Int Psychogeriatr* 25: 185-203.
42. Yury CA, Fisher JE (2007) Meta-analysis of the effectiveness of atypical antipsychotics for the treatment of behavioural problems in persons with dementia. *Psychother Psychosom* 76: 213-218.
43. Lee PE, Gill SS, Freedman M, Bronskill SE, Hillmer MP, et al. (2004) Atypical antipsychotic drugs in the treatment of behavioural and psychological symptoms of dementia: systematic review. *BMJ* 329: 75.
44. Volicer L (2012) Antipsychotics do not have to be used "off label" in dementia. *J Am Med Dir Assoc* 13: 495-496.
45. Ballard C, Creese B, Corbett A, Aarsland D (2011) Atypical antipsychotics for the treatment of behavioral and psychological symptoms in dementia, with a particular focus on longer term outcomes and mortality. *Expert Opin Drug Saf* 10: 35-43.
46. Conn DK, Seitz DP (2010) Advances in the treatment of psychiatric disorders in long-term care homes. *Curr Opin Psychiatry* 23: 516-521.
47. Ballard C, Corbett A, Chitramohan R, Aarsland D (2009) Management of agitation and aggression associated with Alzheimer's disease: controversies and possible solutions. *Curr Opin Psychiatry* 22: 532-540.
48. Burke AD, Tariot PN (2009) Atypical antipsychotics in the elderly: a review of therapeutic trends and clinical outcomes. *Expert Opin Pharmacother* 10: 2407-2414.
49. González-Concepción JJ, Geil K, Jiménez-Velázquez IZ, Ramos-Romey C (2009) Atypical antipsychotics in the management of behavioral symptoms associated with dementia. *Bol Asoc Med P R* 101: 51-53.
50. Ballard CG, Gauthier S, Cummings JL, Brodaty H, Grossberg GT, et al. (2009) Management of agitation and aggression associated with Alzheimer disease. *Nat Rev Neurol* 5: 245-255.
51. Salzman C, Jeste DV, Meyer RE, Cohen-Mansfield J, Cummings J, et al. (2008) Elderly patients with dementia-related symptoms of severe agitation and aggression: consensus statement on treatment options, clinical trials methodology, and policy. *J Clin Psychiatry* 69: 889-898.
52. Buhr GT, White HK (2007) Difficult behaviors in long-term care patients with dementia. *J Am Med Dir Assoc* 8 (3 Suppl 2): e101-e113.
53. Buhr GT, White HK (2006) Difficult behaviors in long-term care patients with dementia. *J Am Med Dir Assoc* 7: 180-192.
54. Aupperle P (2006) Management of aggression, agitation, and psychosis in dementia: focus on atypical antipsychotics. *Am J Alzheimers Dis Other Dement* 21: 101-108.
55. Sutor B, Rummans TA, Smith GE (2001) Assessment and management of behavioral disturbances in nursing home patients with dementia. *Mayo Clin Proc* 76: 540-550.
56. Salzman C, Jeste DV, Meyer RE, Cohen-Mansfield J, Cummings J, et al. (2008) Elderly patients with dementia-related symptoms of severe agitation and aggression: consensus statement on treatment options, clinical trials methodology, and policy. *J Clin Psychiatry* 69: 889-898.
57. Jeste DV, Blazer D, Casey D, Meeks T, Salzman C, et al. (2008) ACNP White Paper: update on use of antipsychotic drugs in elderly persons with dementia. *Neuropsychopharmacology* 33: 957-970.
58. Locca JF, Büla CJ, Zumbach S, Bugnon O (2008) Pharmacological treatment of behavioral and psychological symptoms of dementia (BPSD) in nursing homes: development of practice recommendations in a Swiss canton. *J Am Med Dir Assoc* 9: 439-448.
59. Snowden M, Sato K, Roy-Byrne P (2003) Assessment and treatment of nursing home residents with depression or behavioral symptoms associated with dementia: a review of the literature. *J Am Geriatr Soc* 51: 1305-1317.
60. Azermai M, Petrovic M, Elseviers MM, Bourgeois J, Van Bortel LM, et al. (2012) Systematic appraisal of dementia guidelines for the management of behavioural and psychological symptoms. *Ageing Res Rev* 11: 78-86.

61. Higgins JPT, Green S (2011) *Cochrane handbook for systematic reviews of interventions*. The Cochrane Collaboration. West Sussex, England.
62. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, et al. (2009) The PRISMA statement for reporting systematic reviews and meta-analysis of studies that evaluate health care interventions: explanation and elaboration. *Ann Intern Med* 151: w65-w94.
63. Cheng HY (2009) Assessing the quality of evidence from randomized, controlled drug and nutritional supplement trials conducted among nursing home residents between 1968 and 2004: what can we learn? *J Am Med Dir Assoc* 10: 28-35.
64. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, et al. (2010) CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomized trials. *J Clin Epidemiol* 63: e1-e37.
65. Gross CP, Mallory R, Heiat A, Krumholz HM (2002) Reporting the recruitment process in clinical trials: who are these patients and how did they get there? *Ann Intern Med* 137: 10-16.
66. Meissner K, Fässler M, Rücker G, Kleijnen J, Hróbjartsson A, et al. (2013) Differential effectiveness of placebo treatments: a systematic review of migraine prophylaxis. *JAMA Intern Med* 173: 1941-1951.
67. Barrett B, Brown R, Rakel D, Rabago D, Marchand L, et al. (2011) Placebo effects and the common cold: a randomized controlled trial. *Ann Fam Med* 9: 312-322.
68. Wampold BE, Imel ZE, Minami T (2007) The story of placebo effects in medicine: evidence in context. *J Clin Psychol* 63: 379-390.
69. Wampold BE, Imel ZE, Minami T (2007) The placebo effect: "relatively large" and "robust" enough to survive another assault. *J Clin Psychol* 63: 401-403.
70. Miller FG, Rosenstein DL (2006) The nature and power of the placebo effect. *J Clin Epidemiol* 59: 331-335.
71. Walsh BT, Sysko R (2005) Placebo control groups in trials of major depressive disorder among older patients. *J Clin Psychopharmacol* 25 (4 Suppl 1): S29-S33.
72. Wampold BE, Minami T, Tierney SC, Baskin TW, Bhati KS (2005) The placebo is powerful: estimating placebo effects in medicine and psychotherapy from randomized clinical trials. *J Clin Psychol* 61: 835-854.
73. Streim JE, Porsteinsson AP, Breder CD, Swanink R, Marcus R, et al. (2008) A randomized, double-blind, placebo-controlled study of aripiprazole for the treatment of psychosis in nursing home patients with Alzheimer disease. *Am J Geriatr Psychiatry* 16: 537-550.
74. Mintzer JE, Tune LE, Breder CD, Swanink R, Marcus RN, et al. (2007) Aripiprazole for the treatment of psychoses in institutionalized patients with Alzheimer dementia: a multicenter, randomized, double-blind, placebo-controlled assessment of three fixed doses. *Am J Geriatr Psychiatry* 15: 918-931.
75. Zhong KX, Tariot PN, Mintzer J, Minkwitz MC, Devine NA (2007) Quetiapine to treat agitation in dementia: a randomized, double-blind, placebo-controlled study. *Curr Alzheimer Res* 4: 81-93.
76. Mintzer J, Greenspan A, Caers I, Van Hove I, Kushner S, et al. (2006) Risperidone in the treatment of psychosis of Alzheimer disease: results from a prospective clinical trial. *Am J Geriatr Psychiatry* 14: 280-291.
77. Tariot PN, Schneider L, Katz IR, Mintzer JE, Street J, et al. Quetiapine treatment of psychosis associated with dementia: a double-blind, randomized, placebo-controlled clinical trial. *Am J Geriatr Psychiatry* 14: 767-776.
78. Ballard C, Margallo-Lana M, Juszczak E, Douglas S, Swann A, et al. (2005) Quetiapine and rivastigmine and cognitive decline in Alzheimer's disease: randomised double blind placebo controlled trial. *BMJ* 330: 874.
79. Brodaty H, Ames D, Snowdon J, Woodward M, Kirwan J, et al. (2003) A randomized placebo-controlled trial of risperidone for the treatment of aggression, agitation, and psychosis of dementia. *J Clin Psychiatry* 64: 134-143.
80. Street JS, Clark WS, Gannon KS, Cummings JL, Bymaster FP, et al. (2000) Olanzapine treatment of psychotic and behavioral symptoms in patients with Alzheimer disease in nursing care facilities: a double-blind, randomized, placebo-controlled trial. The HGEU Study Group. *Arch Gen Psychiatry* 57: 968-976.
81. De Deyn PP, Rabheru K, Rasmussen A, Bocksberger JP, Dautzenberg PL, et al. (1999) A randomized trial of risperidone, placebo, and haloperidol for behavioral symptoms of dementia. *Neurology* 53: 946-955.
82. Deberdt WG, Dysken MW, Rappaport SA, Feldman PD, Young CA, et al. (2005) Comparison of olanzapine and risperidone in the treatment of psychosis and associated behavioral disturbances in patients with dementia. *Am J Geriatr Psychiatry* 13: 722-730.
83. De Deyn PP, Carrasco MM, Deberdt W, Jeandel C, Hay DP, et al. (2004) Olanzapine versus placebo in the treatment of psychosis with or without associated behavioral disturbances in patients with Alzheimer's disease. *Int J Geriatr Psychiatry* 19: 115-126.
84. Katz IR, Jeste DV, Mintzer JE, Clyde C, Napolitano J, et al. (1999) Comparison of risperidone and placebo for psychosis and behavioral disturbances associated with dementia: a randomized, double-blind trial. Risperidone Study Group. *J Clin Psychiatry* 60: 107-115.
85. Moore KL, Boscardin WJ, Steinman MA, Schwartz JB (2012) Age and sex variation in prevalence of chronic medical conditions in older residents of U.S. nursing homes. *J Am Geriatr Soc* 60: 756-764.
86. Onder G, Liperoti R, Fialova D, Topinkova E, Tosato M, et al. (2012) Polypharmacy in nursing home in Europe: results from the SHELTER study. *J Gerontol A Biol Sci Med Sci* 67: 698-704.
87. Vetrano DL, Tosato M, Colloca G, Topinkova E, Fialova D, et al. (2013) Polypharmacy in nursing home residents with severe cognitive impairment: results from the SHELTER Study. *Alzheimers Dement* 9: 587-593.
88. Dwyer LL, Han B, Woodwell DA, Rechtsteiner EA (2010) Polypharmacy in nursing home residents in the United States: results of the 2004 National Nursing Home Survey. *Am J Geriatr Pharmacother* 8: 63-72.
89. Leon C, Gerretsen P, Uchida H, Suzuki T, Rajji T, et al. (2010) Sensitivity to antipsychotic drugs in older adults. *Curr Psychiatry Rep* 12: 28-33.
90. Carroll CB, Zajicek JP (2011) Designing clinical trials in older people. *Maturitas* 68: 337-341.
91. Fischer L, Knaebel HP, Golcher H, Bruckner T, Diener MK, et al. (2012) To whom do the results of the multicenter, randomized, controlled INSECT trial (ISRCTN 24023541) apply?--assessment of external validity. *BMC Surg* 12: 2.
92. Petersen MK, Andersen KV, Andersen NT, Søballe K (2007) "To whom do the results of this trial apply?" External validity of a randomized controlled trial involving 130 patients scheduled for primary total hip replacement. *Acta Orthop* 78: 12-18.
93. Rothwell PM (2005) External validity of randomised controlled trials: "to whom do the results of this trial apply?" *Lancet* 365: 82-93.
94. Posternak MA, Zimmerman M, Keitner GI, Miller IW (2002) A reevaluation of the exclusion criteria used in antidepressant efficacy trials. *Am J Psychiatry* 159: 191-200.
95. Zimmerman M, Mattia JI, Posternak MA (2002) Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice? *Am J Psychiatry* 159: 469-473.
96. Zetin M, Hoepner CT (2007) Relevance of exclusion criteria in antidepressant clinical trials: a replication study. *J Clin Psychopharmacol* 27: 295-301.
97. Zimmerman M, Chelminski I, Posternak MA (2004) Exclusion criteria used in antidepressant efficacy trials: consistency across studies and representativeness of samples included. *J Nerv Ment Dis* 192: 87-94.
98. Weisberg HI, Hayden VC, Pontes VP (2009) Selection criteria and generalizability within the counterfactual framework: explaining the paradox of antidepressant-induced suicidality? *Clin Trials* 6: 109-118.

99. Cohen-Mansfield J (2008) Agitated behavior in persons with dementia: the relationship between type of behavior, its frequency, and its disruptiveness. *J Psychiatr Res* 43: 64-69.
100. Whall AL, Colling KB, Kolanowski A, Kim H, Son Hong GR, et al. (2008) Factors associated with aggressive behavior among nursing home residents with dementia. *Gerontologist* 48: 721-731.
101. Bartels SJ, Horn SD, Smout RJ, Dums AR, Flaherty E, et al. (2003) Agitation and depression in frail nursing home elderly patients with dementia: treatment characteristics and service use. *Am J Geriatr Psychiatry* 11: 231-238.
102. Godwin M, Ruhland L, Casson I, MacDonald S, Delva D, et al. (2003) Pragmatic controlled clinical trials in primary care: the struggle between external and internal validity. *BMC Med Res Methodol* 3: 28.
103. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, et al. (2008) Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 337: a2390.
104. Hotopf M (2002) The pragmatic randomised controlled trial. *APT* 8: 326-333.
105. Treweek S, Zwarenstein M (2009) Making trials matter: pragmatic and explanatory trials and the problem of applicability. *Trials* 10: 37.
106. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A (2011) GRADE guidelines: a new series of articles in the *Journal of Clinical Epidemiology*. *J Clin Epidemiol* 64: 380-382.
107. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, et al. (2011) GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 64: 383-394.
108. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, et al. (2011) GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol* 64: 395-400.
109. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, et al. (2011) GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 64: 401-406.
110. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, et al. (2011) GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). *J Clin Epidemiol* 64: 407-415.
111. Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, et al. (2011) GRADE guidelines: 5. Rating the quality of evidence--publication bias. *J Clin Epidemiol* 64: 1277-1282.
112. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, et al. (2011) GRADE guidelines 6. Rating the quality of evidence--imprecision. *J Clin Epidemiol* 64: 1283-1293.
113. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, et al. (2011) GRADE guidelines: 7. Rating the quality of evidence--inconsistency. *J Clin Epidemiol* 64: 1294-1302.
114. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, et al. (2011) GRADE guidelines: 8. Rating the quality of evidence--indirectness. *J Clin Epidemiol* 64: 1303-1310.
115. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, et al. (2011) GRADE guidelines: 9. Rating up the quality of evidence. *J Clin Epidemiol* 64: 1311-1316.
116. Owens DK, Lohr KN, Atkins D, Treadwell JR, Reston JT, et al. (2010) AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions--agency for healthcare research and quality and the effective health-care program. *J Clin Epidemiol* 63: 513-523.
117. Berkman ND, Lohr KN, Ansari MT, Balk EM, Kane R, et al. (2014) Grading the strength of a body of evidence when assessing health care interventions: an EPC update. *J Clin Epidemiol*.
118. Mustafa RA, Santesso N, Brozek J, Akl EA, Walter SD, et al. (2013) The GRADE approach is reproducible in assessing the quality of evidence of quantitative evidence syntheses. *J Clin Epidemiol* 66: 736-742.
119. Murad MH, Altayar O, Bennett M, Wei JC, Claus PL, et al. (2014) Using GRADE for evaluating the quality of evidence in hyperbaric oxygen therapy clarifies evidence limitations. *J Clin Epidemiol* 67: 65-72.

Acknowledgements

Thanks Karen Knight for her guidance on searching literature. Thanks Natalie May for recruiting the medical student. Thanks Margaret Plews-Ogan and University of Virginia School of Medicine for financial support for MSSRP for one author (TNH).

Funding

There was no funding to support this systematic review. One of authors (TNH) was supported by the Medical Student Summer Research Program (MSSRP) from School of Medicine at the University of Virginia and a Bristol-Myers Squibb Foundation Grant together on Diabetes grant. However, neither funding agency had any role in the design, data collection and analysis, or manuscript writing of this systematic review.