REVIEW

Medication development for agitation and aggression in Alzheimer disease: review and discussion of recent randomized clinical trial design

Maria Soto,1 Sandrine Andrieu,1,2 Fati Nourhashemi,1 Pierre Jean Ousset,1 Clive Ballard,3 Philippe Robert,4 Bruno Vellas,1 Constantine G. Lyketsos5 and Paul B. Rosenberg5

1 Gerontopôle, INSERM U 1027, Alzheimer’s Disease Research and Clinical Center, Toulouse University Hospital, France
2 Department of Epidemiology, Toulouse University Hospital Toulouse, France
3 Wolfson Centre for Age-Related Diseases, King’s College, London, UK
4 EA CoBEK / ICMRR University of Nice Sophia Antipolis -CHU, France
5 Department of Psychiatry, The Johns Hopkins Bayview Medical Center, Baltimore, MD, USA

ABSTRACT

Background: The management of disruptive neuropsychiatric symptom (NPS) such as agitation and aggression (A/A) is a major priority in caring for people with Alzheimer’s disease (AD). Few effective pharmacological or non-pharmacological options are available. Results of randomized clinical trials (RCTs) of drugs for A/A have been disappointing. This may result from the absence of biological efficacy for medications tested in treating A/A. It may also be related to methodological issues such as the choice of outcomes. The aim of this review was to highlight key methodological issues pertaining to RCTs of current and emerging medications for the treatment of A/A in AD.

Methods: We searched PubMed/Medline, Cochrane Central Register of Controlled Trials, and ClinicalTrials.gov for RCTs comparing medications with either placebo or other drugs in the treatment of A/A in AD, between January 2008 and December 2013.

Results: We identified a total of 18 RCTs; of these, 11 were completed and 7 ongoing. Of the ongoing RCTs, only one is in Phase III. Seven of 10 completed RCTs with reported results did not report greater benefit from drug than placebo. Each of the completed RCTs used a different definition of “clinically significant A/A.” There was considerable heterogeneity in study design. The primary endpoints were largely proxy-based but a variety of scales were used. The definition of caregiver and scales used to assess caregiver outcomes were similarly heterogeneous. Placebo response was notable in all trials.

Conclusions: This review highlights a great heterogeneity in RCTs design of drugs for A/A in AD and some key methodological issues such as definition of A/A, choice of outcome measures and caregiver participation that could be addressed by an expert consensus to optimize future trials design.

Key words: behavior, agitation, aggression, Alzheimer’s, measurement, therapeutics, clinical trial

Introduction

The number of people living with dementia worldwide is estimated at 35.6 million and expected to increase to 115.4 million by 2050 (Prince et al., 2013). Seventy percent of dementia is due to Alzheimer’s disease (AD); 98% of people with AD (PwAD) develop at least one neuropsychiatric symptom (NPS) over the course of the disease (Steinberg et al., 2008; Gonfrier et al., 2012). At least 20% of outpatients (Lyketsos et al., 2000) and 40% of long-term care residents (Selbæk et al., 2013) exhibit disrupted NPS such as agitation and aggression (A/A) encompassing a range of affective, verbal, and motor disturbances such as restlessness, cursing, aggression, hyperactivity, combative ness, wandering, repetitive calling out, irritability, and disinhibition (Cohen-Mansfield et al., 1995). A/A
tends to co-occur with sleep disorders, delusions, hallucinations, anxiety or dysphoria (Canevelli et al., 2013). A/A is associated with greater caregiver burden (Okura and Langa, 2011), earlier institutionalization and death (Okura et al., 2011), poorer functioning (Okura et al., 2010), greater cost of care (Murman et al., 2002), and more acute hospitalizations (Soto et al., 2012). Thus, the management of A/A is a major priority in caring for PwAD.

Consensus guidelines and expert opinion statements recommend non-pharmacological approaches to be first line (Benoit et al., 2006; Lyketsos et al., 2006; Rabins et al., 2007; Gauthier et al., 2010; Kales et al., 2014) but there are limited options. Examples include caregiver education, training in problem solving, and targeted interventions to causes for specific behaviors (Gitlin et al., 2012). Patients in both community dwelling (CD) (Brodaty et al., 2012) and nursing home (NH) settings benefit (Deudon et al., 2009; Husebo et al., 2011; Ritcher et al., 2012).

Pharmacological treatment for A/A is recommended when non-pharmacological interventions fail or when A/A is linked to dangerousness to others or marked distress. The most studied medication class is antipsychotics (APs), both conventional and atypical. Between 1999 and 2008, several RCTs assessed APs for treating A/A in PwAD. Eleven RCTs used conventional APs, which mostly involved small sample sizes and with durations of 4 and 12 weeks (Schneider et al., 1990; De Deyn et al., 1999; Teri et al., 2000; Lonergan et al., 2002; Ballard et al., 2009). Outcome was defined as a 30% improvement on standardized behavioral rating scales, as per convention. A high placebo response was found in these RCTs. Since 1995, 18 RCTs have examined the efficacy of atypical APs in patients with AD, mainly with durations of 6–12 weeks (only three trials of 6–12 months) (Ballard and Howard, 2006; Schneider et al., 2006). Small scale trials of treatment with drugs other than APs (antidepressants and anti-convulsants mood stabilizers) have produced equivocal results (Ballard et al., 2009). The available data are limited by small numbers of subjects or shortcomings in study design such as the (non-random) statistical distribution of behavior test scores and lack of consideration of effect size.

In a general description all the previous studies since 1990 were placebo-controlled and were parallel-group, fixed-dose controlled, or adjustable/titrated-dose trials, in the majority involving nursing home patients with a mean age over 80 years of age. Among subjects studied, there was a wide degree of variation in type and severity of symptomatology. The clinical trials endpoints were based on behavior rating scales, including the Brief Psychiatric Rating Scale (BPRS), the Behavior Pathology in Alzheimer’s Disease Rating Scale (BEHAVE-AD), the Neuropsychiatric Inventory (NPI), the Cohen-Mansfield Agitation Inventory (CMAI) and subscales (proxy-based more common than direct observation), and global assessments (Salzman et al., 2008). A non-pharmacologic intervention before enrolling a patient in a clinical trial and a placebo run-in period were not common. Repeated measurement analyses were not performed in most trials.

Atypicals APs, mainly risperidone, have the best evidence for short-term efficacy (6–12 weeks), although meta-analyses have not indicated significant benefit for non-aggressive symptoms of agitation (Ballard and Howard, 2006; Schneider et al., 2006). Efficacy is modest and AP use is associated with serious adverse effects including cerebrovascular accidents and mortality (Schneider et al., 2006; Kales et al., 2012; Langballe et al., 2013). In the European Union, risperidone is indicated for the short-term treatment of severe aggression. In Australia the regulatory authority, the Pharmaceutical Benefits Advisory Committee (PBAC), indicates risperidone for the treatment of psychotic symptoms and aggression with unsuccessful non-pharmacological methods. The Food and Drug Administration (FDA) has published a black box warning for the use of atypical APs in PwAD. In North America there are no approved drugs for treatment of NPS in AD. As a result, most agents are used off-label (Maher et al., 2011). Thus, management of severe, persistent or recurrent A/A unresponsive to non-pharmacologic intervention is a real challenge for clinicians.

Emerging neurobiological research about pathogenesis has led to investigation of repositioned and novel therapeutics for A/A in PwAD, as an alternative to APs. However, the limited benefits reported so far may result from limited understanding of pathogenesis but also from key methodological issues. We hypothesized that a great heterogeneity in the design of recent RCTs of drugs for A/A in AD would be found and that specific key methodological issues could be identified. Thus, the objective on this paper was to review methodological aspects from recent RCTs of drugs for A/A in AD since 2008; the date of the most recent consensus statement on clinical trials methodology of treatments for A/A in dementia (Salzman et al., 2008).

Methods

Reports of RCTs of medications for the treatment of A/A in AD published in the English language were
identified by searching PubMed between January 2008 and December 2013, using terms (“Dementia”[Mesh] OR “Alzheimer”[Mesh]) AND (“Clinical Trial”[Mesh] OR “therapeutics”[MeSH Terms]) AND (“Agitation”[Mesh] OR “Aggression”[Mesh] OR “Behavioral symptoms”[Mesh]). Free text was used to identify articles on “neuropsychiatric symptoms,” “treatment for neuropsychiatric behaviors,” and “behavioral and psychological symptoms of dementia.” This search was supplemented by hand searching of reference lists of selected articles, meta-analyses, and review articles. Google Scholar was searched for additional articles, especially of ongoing RCTs and new drugs.

In addition, we searched the Cochrane Central Register of Controlled Trials and ClinicalTrials.gov between January 2008 and December 2013. We included reports (1) whose publication appeared between 2008 and 2013 and (2) registered in a clinical trial registry (including RCTs with no publications or posted results). We included studies where A/A was the primary or co-primary outcome. Studies, where A/A was a secondary outcome, were not included. Studies focused only on psychosis, depression or apathy in PwAD were excluded. Only randomized, parallel-group, controlled trials comparing medication to placebo or to another medication were included.

Results

We identified 18 RCTs evaluating efficacy of medications for treatment of A/A in AD. Of these, 11 were completed RCTs and seven ongoing. These trials were characterized by a great deal of methodological heterogeneity.

The therapeutic agents divide into: (1) repurposed drugs marketed for other indications (e.g., citalopram, dextromethorphan, delta-9-tetrahydrocannabinol (THC) or prazosin) or (2) new chemical entities not approved for any indication (e.g., mibampator or scyllo-inositol). Of these drugs only two were APs.

Table 1 lists RCTs completed between 2008 and 2013. Nine compared drug to placebo; two used an active drug comparator (risperidone vs. escitalopram or risperidone vs. topiramate). Five RCTs had < 50 subjects.

Table 2 lists seven ongoing RCTs, three assessing new chemical entities, and four repurposing drugs marketed for other indications. Only brexipiprazole is in phase III. Most are U.S. trials with some in Europe, Canada, and other regions. The mix of outcomes and industry sponsorship is similar to completed trials.

Completed RCTs do not report superiority of any drug over placebo, or over active comparator, but rather improvement in both groups (Table 1), with three exceptions. The prazosin pilot study (Wang et al., 2009) reported superiority of drug to placebo; the drug is being further studied in an ongoing RCT. Improvements over placebo were reported in a trial of intramuscular aripiprazole (Rappaport et al., 2009). In the very recent CitAD trial, citalopram showed significant improvement compared to placebo on both primary outcome measures (Porsteinsson et al., 2014).

Two trials report no effect on primary outcomes but improvement in secondary outcomes. Mibampator (Trzepacz et al., 2013a) led to better outcomes on the Frontal Systems Behaviors Scale, and memantine-NH on NPI total score (Fox et al., 2012).

The next section highlights key methodological aspects of the completed and on-going RCTs.

Population studied

AGE

Most trials included patients aged between 50 and 90 years; with the youngest of 40 in THC and the oldest of 95 in topiramate versus risperidone trial. There was no age limit in both prazosin RCTs.

DEMENTIA DIAGNOSIS

Diagnosis of AD was based on DSM IV and/or NINDS-ADRDA criteria (McKhann et al., 1984). The ongoing scyllo-inositol trial used recent AD criteria of the National Institute of Aging-AD Association (McKhann et al., 2011). The ongoing THC trial includes patients with vascular or mixed dementia.

DEMENTIA SEVERITY

Dementia severity was assessed on Mini-Mental State Examination (MMSE) (Folstein, 1975) or Clinical Dementia Rating scale (CDR) (Rosen, 1984). The divalproex trial included patients with moderate to severe dementia, while memantine and oxcarbazepine trials studied severe dementia. Citalopram, mibampator, and aripiprazole trials included a wide range of severity. Among ongoing trials, Scyllo-inositol include moderate to moderately-severe AD (MMSE 10–20), while the THC, dextromethorphan-quinidine, and brexipiprazole trials include patients at all severities.

SETTING

Of the 11 completed RCTs, 5 were NH studies, 3 CD studies, while 1 (risperidone vs. escitalopram) assessed hospitalized patients (see Table 1). Only two RCTs (CitAD and prazosin) included patients
<table>
<thead>
<tr>
<th>MOLECULE</th>
<th>MECHANISM OF ACTION</th>
<th>REFERENCE</th>
<th>DESIGN TT PERIOD</th>
<th>PRIMARY AIM DEFINITION</th>
<th>INCLUSION CRITERIA</th>
<th>EXCLUSION CRITERIA</th>
<th>OUTCOME PRIMARY MEASURES</th>
<th>DEFINITION OF PRIMARY COMPARISON &amp; ANALYSIS</th>
<th>SECONDARY OUTCOMES</th>
<th>STUDY LOCATION &amp; SITES</th>
<th>START DATE</th>
<th>END DATE</th>
<th>N* RESULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citalopram</td>
<td>Selective serotonin reuptake inhibitor (SSRI)</td>
<td>Porteniusen et al., 2014</td>
<td>Ph 3 9 wk fixed doses</td>
<td>Agitation NPI ≥ 4 with ≥2 CD &amp; NH NINCDS-ADRDA MMSE 5–28 Major depression</td>
<td>NBRs modified-ADCS-CGIC</td>
<td>ITT LME model of NBRs</td>
<td>NBRs 3–5 p CGIC 20% NPI, MCI, ADL, MMSE, NPI-distress, cumulative lorazepam dose</td>
<td>USA 8</td>
<td>July 2009</td>
<td>Jan 2013</td>
<td>186</td>
<td>Positive for NBRs &amp; mADCS-CGIC</td>
<td></td>
</tr>
<tr>
<td>Mibampator</td>
<td>AMPA (glutamate receptor) potentiator</td>
<td>Trezepson et al., 2013a</td>
<td>Ph 6 12 wk fixed doses</td>
<td>Agitation NPI ≥ 10 AD RSDV MMSE &gt; 26 &amp; NPI &gt; 10 Delirium Vascular dementia</td>
<td>NPI-4 domain (max score 40)</td>
<td>ITT Likelihood-based, MMRM</td>
<td>Effect size of 0.18 in mean change from baseline to end point CMAI, NPI, CGI-I</td>
<td>USA 16</td>
<td>Feb 2009</td>
<td>June 2011</td>
<td>150</td>
<td>Null, positive only for FrSBe</td>
<td></td>
</tr>
<tr>
<td>Memantine-NH</td>
<td>NMDA agonist</td>
<td>Rappaport et al., 2009</td>
<td>Ph 6 24 wk fixed doses</td>
<td>Delirium NPI ≥ 10 MMSE &lt; 10 Previous memantine usage</td>
<td>PANNS-EC</td>
<td>CMAI at 6 wk</td>
<td>Difference of 0.4P from baseline to end point CMAI at 12 wk, NPI, CGI, MMSE, SIB</td>
<td>UK</td>
<td>Sept 2007</td>
<td>Jan 2010</td>
<td>153</td>
<td>Null, positive for total NPI</td>
<td></td>
</tr>
<tr>
<td>Memantine-CD</td>
<td>NMDA agonist</td>
<td>Fox et al., 2012</td>
<td>Ph 3 24 wk fixed doses</td>
<td>Agitation NMPI ≥ 1 for agitation</td>
<td>CMAI at 24 wk</td>
<td>Total NPI SIB</td>
<td>Effect size of 0.31 in mean change from baseline to end point CMAI at 12 wk, NPI, CGI, MMSE, SIB</td>
<td>Canada 23</td>
<td>Dec 2003</td>
<td>Sept 2010</td>
<td>450</td>
<td>Null for all outcomes</td>
<td></td>
</tr>
<tr>
<td>Escitalopram vs. Risperidone</td>
<td>SSRI vs. atypical antipsychotic</td>
<td>Herrmann et al., 2013</td>
<td>Ph 6 8 wk fixed doses</td>
<td>Delirium</td>
<td>Total NPI</td>
<td>Not reported</td>
<td>Time to discontinuation drug</td>
<td>Israel 1</td>
<td>April 2008</td>
<td>May 2010</td>
<td>40</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Topiramate vs. Risperidone</td>
<td>AMPA/GABA modulator vs. atypical antipsychotic</td>
<td>Mowla et al., 2010</td>
<td>Ph 3 8 wk fixed doses</td>
<td>NPI ≥ 4.5 on 1 domain</td>
<td>CMAI CGI</td>
<td>Total NPI</td>
<td>Not reported</td>
<td>MMSE</td>
<td>Iran 1</td>
<td>Jan 2008</td>
<td>Mar 2009</td>
<td>48</td>
<td>Null</td>
</tr>
</tbody>
</table>
Table 1. Continued.

<table>
<thead>
<tr>
<th>MOLECULE</th>
<th>MECHANISM OF ACTION</th>
<th>REFERENCE</th>
<th>DESIGN TT PERIOD</th>
<th>PRIMARY AIM DEFINITION</th>
<th>INCLUSION CRITERIA</th>
<th>EXCLUSION CRITERIA</th>
<th>OUTCOME PRIMARY MEASURES</th>
<th>DEFINITION OF PRIMARY COMPARISON &amp; ANALYSIS</th>
<th>DESIRED CLINICAL DIFFERENCE</th>
<th>SECONDARY OUTCOMES</th>
<th>STUDY LOCATION &amp; SITES</th>
<th>START DATE</th>
<th>END DATE</th>
<th>N°</th>
<th>RESULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valproate</td>
<td>Enhanced GABAergic function</td>
<td>Tariot et al., 2011</td>
<td>24 month fixed doses</td>
<td>Agitation &amp; psychosis ≥ 3 on 1 or more NPI items A/A, D, H for 2 wks</td>
<td>CD, ≥54 y MMSE10–20 NINDS-ADRDA NPI &lt; 1 in A/A, D, H since AD onset</td>
<td>Time to ≥ 3 on 1 or more NPI items A/A, D, H for 2 wks</td>
<td>Time to endpoint Cox proportional hazard ratio.</td>
<td>33% difference in incidence rate</td>
<td>Total NPI Total CMAI ADCS-CGHC Qol-AD MMSE ADAS-cog ADCS-ADL CDR</td>
<td>USA 46</td>
<td>Nov 2005</td>
<td>Mar 2009</td>
<td>313</td>
<td>Null for all outcomes, significant adverse events</td>
<td></td>
</tr>
<tr>
<td>Prazosin</td>
<td>α1-adrenoceptor antagonist</td>
<td>Wang et al., 2009</td>
<td>Pilot flexible doses</td>
<td>Agitation &amp; Aggression ≥ 2 wk BPRS ≥ 4 in at least 1 item</td>
<td>CD &amp; NH No age limit NINDS-ADRDA hypotension</td>
<td>Total NPI CGIC LME model (NPLBPRS) Mann-Whitney test (CGHC)</td>
<td>Blood pressure</td>
<td>BPRS SE</td>
<td>USA 1</td>
<td>Jan 2005</td>
<td>Sep 2007</td>
<td>22</td>
<td>Positive for all outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Melatonin</td>
<td>Melatonin receptor agonist</td>
<td>German et al., 2009</td>
<td>10 days fixed doses</td>
<td>Agitation &amp; Sleep NH ≥ 60 y NINDS-ADRDA Severe AD</td>
<td>Sleep: actigraphy Total sleep time. Analysis of variances</td>
<td>Not reported</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>USA 1</td>
<td></td>
<td>43</td>
<td>Null for all outcomes</td>
<td></td>
</tr>
<tr>
<td>Oxcarbazepine</td>
<td>GABA receptor antagonist</td>
<td>Sommer et al., 2009</td>
<td>Ph 3 8 wk fixed doses</td>
<td>Agitation &amp; Aggression NPI A/A ≥ 6</td>
<td>NH ≥ 55 y Vascular or AD; ICD-10 MMSE 5–16</td>
<td>Changes in Psychotropic drugs</td>
<td>NPI-NH A/A repeated-measures model with an autoregressive covariance structure</td>
<td>difference between treatments of 1.2 p</td>
<td>BARS, NPI-caregiver distress,</td>
<td>Norway 35</td>
<td>Sept 2005</td>
<td>Oct 2006</td>
<td>103</td>
<td>Null for all outcomes</td>
<td></td>
</tr>
</tbody>
</table>

* Completed RCT registered in a clinical trial registry between 2008 and 2013 and/or published between 2008 and 2013.

b The NPI-4 A/A is a 4-domain subscale chosen which combines from the NPI-10 four domains: agitation/aggression; aberrant motor behavior; irritability/emotional lability; and disinhibition.

c This study was prematurely terminated due to recruitment problems.

Abbreviations: Ph = phase; TT = Treatment; y = years; wk = week; NPS = neuropsychiatric symptoms; A/A = agitation/aggression; D/H = delusions/hallucinations; AMB = Aberrant Motor Behavior; s = severity; f = frequency; AAN = American Academy of Neurology; CD = community-dwelling; NH = nursing home; ITT = intention to treat; NINDS-ADRDA = National Institute Of Neurological and Communicative Disorders and Stroke and the Alzheimer’s Disease and Related Disorders Association; DSM = Diagnostic and Statistical Manual of Mental Disorders; NPI = Neuropsychiatric Inventory; GMAI = Cohen-Mansfield Agitation Inventory; ADCS-ADL = Alzheimer’s Disease Cooperative Study-Activities of Daily Living Inventory; ADCS-CGIC = Alzheimer’s Disease Cooperative Study-Clinical Global Impression of Change; FrSeB = FrONTAL Systems Behaviors Scale; CGI-C = Clinical Global Impression of Change; CGI-I = Clinical Global Impression Improvement Score; CGI-S = Clinical Global Impression Severity of Illness Score; PANSS-EC = Positive and Negative Syndrome Scale – Excited Component; ACES = Agitation-Calmness Evaluation Scale; BPRS = Brief Psychiatric Rating Scale; SIB = Severe impairment Battery; CIBIC-plus = Clinician’s Interview-Based Impression of Change-plus version; ABRS = Agitation Behavior Rating Scale; NIA = National Institute of Aging; QoL-AD = Quality of Life-AD; p = points; IN = intramuscular; LME model = linear mixed effects model; ANCOVA = analysis of covariance; MMRM = mixed-effects model repeated measures analysis.
### Table 2. Ongoing randomized controlled trials (RCT) of drugs for agitation and aggression in Alzheimer’s disease

<table>
<thead>
<tr>
<th>MOLECULE</th>
<th>MECHANISM OF ACTION</th>
<th>DESIGN</th>
<th>TT PERIOD</th>
<th>PRIMARY AIM A/A DEFINITION</th>
<th>INCLUSION CRITERIA</th>
<th>EXCLUSION CRITERIA</th>
<th>OUTCOME PRIMARY MEASURES</th>
<th>SECONDARY OUTCOMES</th>
<th>STUDY LOCATION &amp; SITES</th>
<th>START DATE</th>
<th>END DATE</th>
<th>N (TARGET)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scyllo-inositol (ELND005)</td>
<td>Myo-inositol metabolism &amp; phosphoinositol signaling Anti-amyloid</td>
<td>Ph 2</td>
<td>12 wk fixed doses</td>
<td>Agitation &amp; Aggression NPI A/A ≥ 4 (≥ 2 &amp; ≥ 2)</td>
<td>CD 50–88 y NIA-AA MMSE 10–21</td>
<td>Major depression Psychosis</td>
<td>NPI-C combined A/A scores</td>
<td>ADCS-CGIC, NPI, NPI-C domains, CMAI, ADCS-ADL, MMSE</td>
<td>USA Canada Europe</td>
<td>Nov 2012</td>
<td>Dec 2013</td>
<td>400</td>
</tr>
<tr>
<td>Prazosin</td>
<td>α1-adrenoreceptor antagonist</td>
<td>Ph 2</td>
<td>12 wk fixed doses 12 wk open label</td>
<td>Disrupted agitated behavior at least twice/wk</td>
<td>CD No age limit</td>
<td>Hypotension</td>
<td>ADCS-CGIC &amp; total NPI</td>
<td>Number of days completed BPRS Total BPRS</td>
<td>USA 1</td>
<td>March 2010</td>
<td>July 2015</td>
<td>120</td>
</tr>
<tr>
<td>Brexpiprazole (OPC-34712)</td>
<td>Dopamine D2 receptor partial agonist</td>
<td>Ph 3</td>
<td>12 wk 3 fixed Doses 30 days follow-up</td>
<td>Agitation NPI-A/A ≥ 4</td>
<td>NH 55–90 y NINDS-ADRDA MMSE 5–22</td>
<td>Axis I (DSM 4) Axis II</td>
<td>Total CMAI CGI-S, total NPI NH, CMAI subscales, NPI individual items, NPI-D/H, caregiver distress, CGI-I, CGI-E, QoL-AD</td>
<td>USA, Canada, Europe</td>
<td>July 2013</td>
<td>June 2017</td>
<td>560</td>
<td></td>
</tr>
<tr>
<td>Brexpiprazole (OPC-34712)</td>
<td>Dopamine D2 receptor partial agonist</td>
<td>Ph 3</td>
<td>12 wk flexible Doses 30 days follow-up</td>
<td>Agitation CGI-S ≥ 4</td>
<td>NH 55–90 y NINDS-ADRDA MMSE 5–22</td>
<td>Axis I (DSM 4) Axis II</td>
<td>Total CMAI CGI-S, total NPI NH, CMAI subscales, NPI individual items, NPI-D/H, caregiver distress, CGI-I, CGI-E, QoL-AD</td>
<td>USA, Canada, Europe</td>
<td>Sept 2013</td>
<td>May 2016</td>
<td>230</td>
<td></td>
</tr>
<tr>
<td>Delta-9-tetrahydrocannabinol</td>
<td>CB-1 receptor agonist</td>
<td>Pilot</td>
<td>Ph2</td>
<td>2 wk fixed dose</td>
<td>NPS NPI ≥ 10 and at least 1 domain A/A or AMB ≥ 1</td>
<td>CD VaD, AD or AD/VaD ≥ 18 y CDR 1–3 Persistent pain</td>
<td>Major psychiatric disorder</td>
<td>Total NPI CMAI, Zarit.</td>
<td>Netherlands 2</td>
<td>Sept 2011</td>
<td>Dec 2013</td>
<td>22</td>
</tr>
<tr>
<td>Delta-9-tetrahydrocannabinol</td>
<td>CB-1 receptor agonist</td>
<td>Ph2</td>
<td>3 wk fixed dose</td>
<td>NPS NPI ≥ 10 and at least 1 domain A/A or AMB ≥ 1</td>
<td>CD VaD, AD or AD/VaD ≥ 40 y CDR 1–3 Persistent pain</td>
<td>Major psychiatric disorder</td>
<td>Total NPI CMAI, PACSLAC-D, Caregiver-CGIC, QoL-AD, Barthel</td>
<td>Netherlands 2</td>
<td>June 2012</td>
<td>Feb 2014</td>
<td>150</td>
<td></td>
</tr>
</tbody>
</table>
## Table 2. Continued.

<table>
<thead>
<tr>
<th>STUDY LOCATION &amp; SITES</th>
<th>PRIMARY OUTCOME MEASURES</th>
<th>PRIMARY AIM A/A DEFINITION</th>
<th>SECONDARY OUTCOME MEASURES</th>
<th>OUTCOME PRIMARY CRITERIA</th>
<th>MECHANISM OF ACTION</th>
<th>MOLECULE</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>ADCS-CGIC, QoL-AD, ADSC-ADL, CSI</td>
<td>Total NPI, ADCS-CGIC, QoL-AD, ADSC-ADL</td>
<td>Dextromethorphan, (AVP-923)</td>
<td>NMDA receptor antagonist and antagonist</td>
<td>Dextromethorphan/Quinidine</td>
<td>Dextromethorphan/Quinidine</td>
</tr>
<tr>
<td>USA</td>
<td>ADCS-CGIC, QoL-AD, ADSC-ADL, CSI</td>
<td>Total NPI, ADCS-CGIC, QoL-AD, ADSC-ADL</td>
<td>NMDA receptor antagonist and antagonist</td>
<td>Dextromethorphan/Quinidine</td>
<td>Dextromethorphan/Quinidine</td>
<td>Dextromethorphan/Quinidine</td>
</tr>
</tbody>
</table>

**Diagnosis of Clinically Significant A/A**

There is not a “gold standard” definition of A/A in AD (Salzman et al., 2008). In its absence two approaches have been used: (1) judgment by experienced clinicians that medication is appropriate and/or (2) severity rating above a cut-off indicative of at least moderate A/A. CitAD and dextromethorphan-quinidine trials combined both. A/A encompass a range of recognizable features (such as hitting, pacing, or disinhibition) but these are shared with other behavioral disturbances of AD. It is not surprising that inclusion criteria vary between RCTs. Several trials have used cutoffs on total NPI (mibampator, and THC). The THC trials additionally require high scores on two domains of the NPI: A/A or motor disturbance.

Although NPI is widely used to define A/A across RCTs, the specific definitions have varied (Tables 1 and 2). Some trials (CitAD, brexpiprazole, and scylo-inositol) used the NPI A/A domain, while the mibampator trial required the presence of at least one of four NPI domains.

Thresholds for agitation severity similarly vary: in memantine-NH trial agitation was high (CMAI > 45) (Fox et al., 2012), while in the memantine CD trial it was low (NPI A/A ≥ 1) (Herrmann et al., 2013). Recent RCTs (CitAD and mibampator), and the ongoing scylo-inositol and brexpiprazole trials use a threshold of ≥4 on NPI A/A; CITAD adds the requirement that A/A occur at least several times/week (“frequently”) to improve relevance and applicability to real-life situations.

Moreover, the CITAD, brexpiprazole, and scylo-inositol trials require a history of poor response to non-pharmacological interventions for inclusion.

The FDA appears to be increasingly accepting A/A as a target indication, examples being the ongoing dextromethorphan and scylo-inositol trials.

### Caregiver Participation

Most trials require a reliable caregiver as informant for patient symptoms and caregiver burden. In CITAD the requirement was for a primary caregiver who spends at least several hours a week with the patient, supervises his/her care and attends clinic visits with the patient. Mibampator required “frequent or daily contact with the patient.” Dextromethorphan requires a caregiver who is with the patient a minimum of three times per week on three separate days, THC requires a caregiver who
Concomitant Psychotropic Medication

All trials allow AD specific medications (cholinesterase inhibitors and memantine) with doses stable for at least one month prior to enrollment. Criteria for concomitant psychotropics vary considerably. Mibamptor and CitAD excluded most concomitant psychotropics (except for “rescue” medications). Both memantine trials allowed psychotropic medications at baseline (including antidepressants, antipsychotics, or anxiolytics), while the ongoing prazosin trial allows such medication if the patient is a partial responder. The ongoing THC trial allows psychotropic medications except tricyclics, fluoxetine or carbamazepine.

Study Design

Randomization

Stratification of randomization varies with some trials stratifying by clinical site and others adding presence of antide pressant treatment (scyllo-inositol), or severity (scyllo-inositol, mibamptor).

Pharmacological Intervention

These are fixed-dose or adjustable/titrated-dose trials. All studies used oral formulations except one trial of intramuscular aripiprazole (Rappaport et al., 2009). In completed RCTs, the duration of treatment was 8–12 weeks except for melatonin (ten days), memantine (24 weeks), or divalproex (24 months). Among ongoing RCTs, the duration of treatment is 12 weeks except for dextromethorphan (10 weeks), and 3 weeks for THC.

Allowed Rescue Medication

All RCTs allowing “rescue” medication for severe A/A used lorazepam with doses ranging from 0.5 mg to 4 mg daily. Trazodone was used at low doses (50 mg per day) for sleep disorders (CitAD) and at higher doses (50–150 mg per day) for severe A/A in the memantine trial (Fox et al., 2012). Haloperidol was used in the oxcarbazepine trial.

Non-pharmacological Intervention

In CitAD, mibamptor and divalproex trials, different interventions, at inclusion and regularly over treatment period, have been used but vary in concept, content, and intensity. In CitAD the psychosocial intervention during the study was more intense than mibamptor, for example. The goal of the CitAD intervention is to systematize education and support in RCTs (Rosenberg et al., 2010) by providing “enhanced usual care” to patients and caregivers regardless of treatment assignment (Drye et al., 2012).

Outcome Measures

NPS measures may be narrow spectrum, (assessing depression or agitation, for example), or broad spectrum (Steinberg and Lyketsos, 2008), including the NPI. The NPI originally included 10 domains (Cummings et al., 1994) but expanded to 12 by the addition of sleep and appetite domains (Cummings et al., 1997). Domains of NPI are commonly used in studies with varied degrees of validation. Denneyh et al. (Denneyh et al., 2012) evaluated a cluster of items from the NPI in the hope of validating A/A assessment. Based on epidemiology and consultation with clinicians, they selected the individual NPI-10 domains of agitation, irritability, disinhibition, and aberrant motor behavior as a four-item measure of agitation and aggression (NPI-4-A/A). The 4-A/A subscale of NPI-Q was validated in the Alzheimer’s Disease Neuroimaging Initiative and National Alzheimer’s Coordinating Center cohorts (Trzepacz et al., 2013b).

One limitation of NPI is that assessment relies solely on subjective caregiver input with resultant bias. Recently, the Neuropsychiatric Inventory Clinician Rating (NPI-C) was developed in part to address this issue (de Medeiros et al., 2010). Aberrant vocalization was added as a new domain, whereas the A/A domain of the NPI was split to arrive at a total of 14 domains. Unlike NPI, each domain and each item within a domain can be rated on the NPI-C. Trained clinicians use input from caregiver and patient to rate frequency, severity, and distress of each item. NPI-C was field tested in an international validation study and compared with focused scales to determine convergent validity (de Medeiros et al., 2010). It was found to have greater reliability than the conventional NPI.

The CMAI (Cohen-Mansfield, 1996) is widely used particularly in more advanced dementia and NH settings. Items are rated on a seven-point scale assessing the frequency of agitated behaviors (ranging from “never” to “several times an hour”). Items are presented in four subscales: aggressive behavior; physical non-aggressive behavior; verbally agitated behavior; hiding and hoarding. There are versions for both settings, NH (29 items) and a non-validated CD (36 items). CMAI provides the richest description of A/A but has the limitations of a relatively long administration time and assessments based solely on subjective caregiver input.

The Alzheimer’s Disease Cooperative Study-Clinical Global Impression of Change (ADCS-CGIC) is a clinician-rated judgment of change
from baseline (Schneider et al., 1997) and is based on the Clinical Global Impression of Change (CGI-C) (Guy, 1976). The modified ADCS-CGIC (mADCS-CGIC) version targets global functioning in specific NPS domains is particularly suited to blinded assessments in RCTs. The CitAD and scyllo-inositol trials have chosen a modified version of the ADCS-CGIC assessing specific A/A related domains.

The divalproex trial (Tariot et al., 2011) uniquely measured incident A/A as the primary outcome. Of 18 RCTs, NPI was the primary outcome in ten trials and a secondary outcome in eight. The subscale NPI-4-A/A domain was primary outcome in the mibampator trial (Trzepacz et al., 2013a). The ongoing scyllo-inositol trial has chosen as primary outcome the NPI-C (combined scores on agitation + aggression domains) and other NPI-C domains as secondary outcomes. Three completed and two ongoing trials, such as the brexpiprazole trials, used and are using CMAI as primary outcome. In eight RCTs CMAI was a secondary outcome. All RCTs used a global impression of change measure as secondary outcome except for both prazosin trials and the ongoing CitAD where they were co-primary outcomes.

Other outcomes include: Alzheimer’s Disease Cooperative Study Activities of Daily Living (ADCS-ADL) (Galasko et al., 1997) in five RCTs and seven quality of life (Qol-AD, Logsdon et al., 2002) in five RCTs. MMSE is the most widely used cognitive (secondary) outcome. Caregiver burden was assessed in four completed and three ongoing trials, using the NPI-caregiver distress item (total or specific items); one THC trial uses the Zarit Burden Inventory (Zarit et al., 1980).

Safety outcomes
Most RCTs of A/A include a typical safety evaluation: adverse events (AE), serious adverse events (SAV), and AEs leading to withdrawal. Specific safety outcomes related to psychotropic medication effects have included cognition (based on MMSE), falls, sedation, weight change, and QT prolongation. After the FDA issued an advisory regarding dose-dependent risk of QT prolongation with citalopram, CITAD increased surveillance of QT intervals and reported notable QT prolongation with citalopram at 30 mg. Balance, gait, and mobility are targets of assessments for CitAD, brexpiprazole, and THC. Brexpiprazole also assesses specific AEs related to AP use such as extra-pyramidal symptoms (Simpson-Angus Scale), tardive dyskinesia (Abnormal Involuntary Movement Scale), and akathisia (Barnes Akathisia Rating Scale).

Most of trials reported AE and SAV based on caregivers and physicians. AE could be recorded as spontaneous but also a checklist could be used like in CitAD and mibampator trials.

Analytic strategies
All statistical analysis plans for these RCTs were intention-to-treat (ITT). Primary comparisons were made using: (1) Analysis of covariance (ANCOVA) with “last observation carried forward” (LOCF) to estimate change from baseline to endpoint in the memantine CD trial, or (2) mixed models (CitAD, mibampator, memantine in NH, and prazosin). Recently, the general superiority of mixed models over LOCF was established (Siddiqui et al., 2009), especially for RCTs with longer treatment periods.

Among mixed models, linear mixed models or mixed models of repeated measures (MMRM) are options. The latter treat time as a continuous or categorical variable, estimating mean change from baseline, adjusting for baseline performance. MMRMs are attractive because they make no assumption about the shape of the outcome’s trajectory over time. A recent paper compared MMRM to linear models using data from several trials (Donohue and Aisen, 2012). Neither approach was more robust to missing data, an important issue in AD trials. CitAD and prazosin used linear effects models with random intercept and slope in the primary comparison, while mibampator used MMRM.

Choice of drug
NPS in AD are thought to reflect one or more types of CNS dysfunction: (1) synaptic or circuit disconnections in specific neuronal networks, (i.e., frontal-subcortical and cortico-cortical networks); (2) dysfunction in ascending monoaminergic systems involving serotonin, norepinephrine, or dopamine neurons primarily located in the brain stem and diffusely projecting via long axons to virtually all parts of the brain; and (3) glutamate-mediated excitatory neurotoxicity. These CNS alterations are not mutually exclusive and likely synergize to mediate NPS.

Scyllo-inositol appears to improve synaptic activity in networks underlying NPS via a dual mechanism of action: (1) regulation of brain myo-inositol metabolism and phosphoinositol signaling, and (2) protection from oligomer-induced toxicity due to amyloid anti-aggregation effects (Townsend et al., 2006). Loss of serotonin in the inferior frontal cortex was reported to be limited to AD patients with prominent aggressive behavior.
Brexipiprazole is chemically similar to aripiprazole and has broad activity across multiple monoamine systems with reduced partial agonism for D2, 5HT1A receptors, and enhanced antagonism for 5-HT2A, and α1-adrenoceptors. There is growing interest in modulators of glutamate neurotransmission, implicated in many neuropsychiatric diseases including schizophrenia and AD (Blanchard et al., 2004). Dextromethorphan hydrobromide (DM) (the active main molecule) modulates glutamate signaling in two ways: (1) pre-synaptic inhibition of glutamate release (by sigma-1 receptor agonism), and (2) post-synaptic glutamate response modulation (by weak blockade of NMDA receptor and modulation of NMDA response to glutamate by the sigma-1 receptor). In the compound, quinidine sulfate increases the bioavailability of DM. DM-quinidine is FDA-approved for treatment of pseudo-bulbar affect whose symptoms overlap partially with A/A in AD. Mibampator is an AMPA glutamatergic receptor potentiator (O’Neill, 2004). Activation of AMPA receptors strengthens synapses and changes in glutamatergic synaptic transmission that contributes to neural plasticity in the central nervous system (Yamada, 2000). Memantine’s mechanism of action is related to glutamatergic synaptic transmission.

Delta-9-tetrahydrocannabinol (THC) is the most biologically active isomer of THC, a psychoactive compound extracted from the resin of Cannabis sativa (marijuana). The ongoing THC trial is testing a very similar molecule to dronabinol, synthetic delta-9-THC, indicated for severe nausea and vomiting caused by chemotherapy, or for anorexia with weight loss in patients with acquired immunodeficiency syndrome. THC activates cannabinoid receptors (mainly type 1), repressing neurotransmitter release in the brain. The rational for its use in NPS is based on its psychoactive effects and association with reduced pain sensation. Although earlier retrospective data reported benefits of dronabinol for A/A in severely demented patients (Woodward et al., 2014), the exact mechanism of action is not known. This could be explained by preventing aggregation of amyloid-beta with consequent microglial activation (Ramirez et al., 2005).

Discussion

We reviewed recent 11 completed and 7 ongoing RCTs of current and emerging medications for the treatment of A/A in AD. The major result of this review is the great heterogeneity found in the design of these RCTs. Only three (pilot prazosin, intramuscular aripiprazole, and CitAD) report results favoring drug treatment of A/A, emphasizing the lack of current evidence for a definitive treatment strategy. Most trials report improvement on placebo likely due to intangible benefits of being in RCTs, variability in outcomes, and variability in the course of A/A (Garre-Olmo et al., 2010), with spontaneous remissions or improvement without treatment. This high rate of response on placebo decreases the statistical power of trials to detect drug effects. Additional challenges include: (1) disentangling treatment effects from competing effects on the outcome in the presence of psychosocial interventions; (2) varied approaches to use of rescue and concomitant psychotropic medications; (3) heterogeneity of target symptoms and overlap of NPS (Lyketsos et al., 2011); (4) variable inclusion criteria; and (5) the use of global AD measures rather than measures specific to A/A or individual A/A symptoms (Drye et al., 2012).

The field does not yet have a tight consensus on primary outcome measures, on what constitutes a clinically significant effect size (Mohlar et al., 2009), or how to translate effect size between different A/A scales. There are two overall approaches to assess treatment response both of which should be used: (1) outcomes based on the judgment of experienced clinicians (such as the mADCS-CGIC) and/or (2) outcomes measuring the severity of A/A symptoms.

The heterogeneous approach toward defining caregiver role and assessing caregiver outcomes is another issue in RCTs of A/A. The field needs consensus on who is a caregiver (family, formal/professional caregiver, time spent with the patient, etc.). Caregiver opinion is clearly necessary in the evaluation of A/A symptoms, which by their nature vary over time, and need to be assessed by history-taking rather than merely observation in the clinic. But caregiver opinion is by its nature subjective and, therefore, vulnerable to reporting bias; there is a role for training the caregiver in the assessment in order to obtain more standardized reports. One important innovation is the NPI-C that provides anchors for experienced clinician judgments that can overcome caregiver biases and allows for input from all available information including the clinical record, caregiver input, and direct observation.
Table 3. Recommendations for future RCTs targeting A/A in patients with AD

<table>
<thead>
<tr>
<th>METHODOLOGICAL ASPECT</th>
<th>RECOMMENDATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population studied</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>No limit</td>
</tr>
<tr>
<td>Dementia severity</td>
<td>Mild to severe based on CDR rating of 1–3; stratification</td>
</tr>
<tr>
<td>Settings</td>
<td>Different RCTs for NH or CD preferred; or stratification</td>
</tr>
<tr>
<td>Clinically significant A/A</td>
<td>A/A needs consensus criteria</td>
</tr>
<tr>
<td></td>
<td>“Clinically significant” = medication is needed based on judgment of experienced clinician combined with severity rating above a cut-off on a A/A scale</td>
</tr>
<tr>
<td>Concomitant medications</td>
<td>“AD treatments” allowed on stable doses for 30–60 days</td>
</tr>
<tr>
<td></td>
<td>APs not allowed; or allowed stable doses for 30–60 days</td>
</tr>
<tr>
<td></td>
<td>Antidepressants, mood stabilizers, anticonvulsants: allowed on stable doses for 30–60 days</td>
</tr>
<tr>
<td>Caregiver participation</td>
<td>Caregiver needs a consensus definition</td>
</tr>
<tr>
<td></td>
<td>Standardized training in recognizing NPS and in rating behavior scales</td>
</tr>
<tr>
<td></td>
<td>Use of a caregiver diary for real time observations</td>
</tr>
<tr>
<td>Study design</td>
<td></td>
</tr>
<tr>
<td>Pharmacological intervention</td>
<td>Run-in-period before randomization (2–4 weeks)</td>
</tr>
<tr>
<td></td>
<td>8–12-week treatment period</td>
</tr>
<tr>
<td></td>
<td>Consolidation response: to assess time to relapse within responders in each group during a 6–12-month period</td>
</tr>
<tr>
<td>Non-pharmacological intervention</td>
<td>Psychosocial intervention during the run-in and the treatment periods in both groups.</td>
</tr>
<tr>
<td></td>
<td>Etiologic, non-pharmacologic, person-centered approach during run-in and treatment periods in both groups</td>
</tr>
<tr>
<td>Allowed rescue medication</td>
<td>Defined allowable dosing, monitored use</td>
</tr>
<tr>
<td>Outcome measures</td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>Global measure of A/A as primary</td>
</tr>
<tr>
<td></td>
<td>Validated scales assessing A/A, co-primary or secondary</td>
</tr>
<tr>
<td></td>
<td>Rated by clinicians with patient and caregiver input</td>
</tr>
<tr>
<td>Secondary</td>
<td>Consider actigraphy</td>
</tr>
<tr>
<td></td>
<td>Agitation symptoms</td>
</tr>
<tr>
<td></td>
<td>Aggression symptoms</td>
</tr>
<tr>
<td></td>
<td>Other NPS: irritability, anxiety, depression, psychosis</td>
</tr>
<tr>
<td></td>
<td>Cognition, functional ability, quality of life</td>
</tr>
<tr>
<td></td>
<td>Caregiver distress, other caregiver measures</td>
</tr>
<tr>
<td></td>
<td>Allowed rescue medication cumulative dose</td>
</tr>
<tr>
<td>Analytic strategies</td>
<td>Intention to treat analysis</td>
</tr>
<tr>
<td></td>
<td>Mixed models: LMM or MMRM</td>
</tr>
</tbody>
</table>

Abbreviations: AD = Alzheimer’s disease; NPS = neuropsychiatric symptoms; A/A = agitation/aggression; CD = community dwelling; NH = nursing home; CDR = clinical dementia rating; MMRM: mixed model of repeated measures; linear mixed models.

Methodological enhancements of recent RCTs: suggestions for improvement

We propose in Table 3 methodological considerations for future trials for A/A in PwAD.

Population studied

DISEASE SEVERITY
The THC trial is the first trial to use the CDR to measure dementia severity for inclusion. The MMSE’s major limitation is that the estimate of disease severity is based solely on cognitive performance, and does not capture, as CDR does, cognitive and physical impairments which are key features of AD. Therefore, disease severity should be based on CDR.

In recent trials there is a tendency to include patients at all levels of disease severity and, thus, patients at milder stages. This is important, given that the pathological processes leading to AD begins decades before a clinical diagnosis, and drugs targeting NPS may need to be assessed earlier in the disease process.

SETTING
The majority of trials have included patients from CD or from NH settings and only two (CitAD and dextromethorphan trials) have patients from
both. It is important for generalizability to include participants from both settings, either in separate trials or in a single trial; the latter strategy might be best implemented by stratifying randomization by care setting (as is being done in CiTAD, for example). Additionally, it is important to validate methods of caregiver data collection since most caregivers in CD are family members and most in NH are paid staff who may have different agendas and different reporting bias. It is important not to exclude NH participants since the majority of A/A in advanced dementia is in the NH setting.

Concomitant Psychotropic Medications

Most patients with A/A will already be taking psychotropic medications, and the approach toward these needs to find a middle ground between maximal rigor (no concomitant medications) and maximal generalizability (allow concomitant medications and assess intervention as an “add-on” therapy). Many RCTs have chosen the former approach to assess efficacy of an intervention as monotherapy, with the trade-offs of rendering recruitment more challenging, possibly inadvertently excluding the most severe A/A, and limiting generalizability. The alternative approach could be proposed for these reasons. The ongoing prazosin trial follows this approach, allowing all stable concomitant psychotropic medications at inclusion, if the patient is a partial responder. However, doses of concomitant psychotropic should remain stable over the treatment period. Regarding APs medication, since they are the only one having proved efficacy, even modest, their concomitant use it is questionable and should be address by an expert panel (see Table 3).

Study design

Placebo run-in

One design feature to diminish the effect of placebo response is to have a placebo run-in period, which is becoming the norm in other fields of psychopharmacology (Iovieno and Papakostas, 2012), accounting for the benefit of non-pharmacologic interventions, and improving statistical power (Frost et al., 2008). None of these trials included a placebo run-in period before randomization. Two trials in related fields used two to four week placebo run-ins and reported relatively low on-placebo response rates in the randomized treatment period (Howard et al., 2007, Cummings et al., 2013). We suggest that for RCTs of A/A the period be kept to two weeks because this is a reasonable period of time for relying solely on non-pharmacologic interventions, the duration needs to be short due to the acuity of symptoms (see Table 3).

Pharmacological Intervention

Most completed and current RCTs administer drug treatment for 9 to 12 weeks, adequate for assessing acute response, but not stability of response over longer periods of time. Some design strategies partly address this issue. The prazosin trial has included a 12-week open label observation period after the strict RCT period. There is an ongoing safety extension RCT to evaluate persistence of the effects of scylo-inositol on A/A, beyond the treatment period covered by the trial. All patients will receive scylo-inositol, but masking from the first trial will be maintained. Another strategy could be a two-week washout after the RCT period to assess if there is a rebound in symptoms.

Open-label observation under treatment drug it could be interesting if the period was long enough to assess “consolidation” of response, preferably 6 to 12 months for maximizing clinical generalizability. However, having no comparator, such a methodology will not inform efficacy. In order to address consolidation efficacy the observation period should be long enough (six to one year) and the analysis should measure time. It will also be interesting to assess time to remission or relapse, as is done in evaluation of anticonvulsants, for example.

Non-pharmacologic intervention

The citalopram and mibampator trials include a psychosocial intervention for all participants regardless of treatment assignment, and this design feature is important to provide ethically “enhanced usual care” to all participants, to improve recruitment and retention, (Drye et al., 2012) and, to reduce variance in outcome reporting by caregivers via education about the features of the disease (Tariot et al., 2011).

Allowed rescue medication

A/A can truly be a crisis for family or institutional caregivers, and rescue medication is often needed to have viable retention. The CiTAD trial is a model, monitoring the use and dose of rescue medication by treatment group (Drye et al., 2012).

Outcome measures

The choice of the optimal primary outcome is probably the most important decision in study design in this field. As this review shows, the choices have been very heterogeneous in completed and ongoing RCTs due to the lack of a gold standard outcome measure. However, the field appears to be
reaching a consensus in using both agitation-specific quantitative measures (e.g., relevant domains of NPI-C or CMAI) plus a global rating of change for agitation outcomes (mADCS-CGIC). This combined approach has been successfully used in the CitAD trial. Although the pimavanserin trial is not included in this review (targeting psychosis in Parkinsons’ disease) the assessment process is worth noting: central raters assessed the primary outcome, site-based raters assessed CGI, and caregivers assessed the Zarit burden scale (Cummings et al., 2013). Assessing different outcomes with different raters may provide robust support for convergent clinical benefits. Actigraphy offers the possibility of objective measurement of activity associated with agitation and has reasonable validation versus NPI in two studies (Mahlberg et al., 2007; Kirste et al., 2014).

Analytical strategies

An issue to take into account in statistical analyses is measurement variance that may be exaggerated in some trials especially multinational studies where translations and cultural interpretations are required. In most recent RCTs (citalopram, mibampator, memantine (Fox et al., 2012) trials) mixed models have been estimated. Currently, mixed models seem to be more appropriated than ANCOVA and LOCF in RCTs of drugs in AD. However, up to date there is no available data to better propose MMRM or linear mixed models as statistical strategy in RCTs of drugs for A/A in AD.

There are several important methodologic obstacles to drug development in this area that might be best amenable to expert consensus on the: (1) definition of A/A for inclusion; (2) choice of primary outcome; (3) role of adaptive designs to minimize exposure to a drug or placebo that is “not working” (Mugno et al., 2004; Kairalla et al., 2012); (4) standardization of non-pharmacological person-centered interventions (Gitlin et al., 2012) in both arms; and (5) placebo run-in-period prior to randomization. Once eligibility is confirmed at baseline, only patients with persisting A/A or at higher levels of A/A severity would participate in the treatment period. This etiologic non-pharmacological centered person approach, in addition to the psychosocial intervention, should be prolonged over the treatment period. This approach will reflect real-life clinical practice, reduce placebo response rate, and address the real-life issue which is whether a drug can demonstrate superior efficacy to placebo in the presence of non-pharmacological intervention for all participants.

Our review has limitations. First, even all studies were randomized and controlled trials and thus with a high quality, differences in the methodological quality were not deeply analyzed. Second, comparisons between RCTs’ reports inter-rater reliability were not available. Third, only publications in English were included.

Finally, in comparison to pre-2008 RCTs, it is notable that other potential pharmacological alternatives to atypical APs, have been and are being tested in recent RCTs. However, despite considerable efforts in crafting appropriate designs for RCTs of promising therapeutics agents for A/A, it is urgent to gain more clarity regarding the underlying neural regions and circuitry involved in NPS, to thereby shed light on symptom pathogenesis. However, so far, there has been very little developmental work based on this better understanding that could offer an opportunity to develop more new targeted drug treatments (Ballard et al., 2013; Geda et al., 2013).

Conclusion

Despite the urgent need to identify effective pharmacological treatments for A/A in PwAD, progress has been slow. In the past six years a small number of RCTs of drugs for treatment of A/A in AD have been conducted with disappointing results. These trials are characterized by methodological heterogeneity. Several issues have been encountered: the need for stronger consensus on the syndromal definition of A/A as a target, choice of primary efficacy outcome measure, the content and timing of the non-pharmacological intervention in placebo and drug arms before and/or during the trial, concomitant psychotropic medication, definition of caregivers and their participation. Consensus is necessary to enhance the design of future trials. The fact that placebo effects are substantial and consistently observed suggests that non-pharmacological approaches are currently the standard of care that we are still waiting for clear and consistent evidence on drug efficacy.

Conflict of interest

Dr Soto has served as consultant for Ethypharm. Dr. Rosenberg has received research support from National Institute on Aging (1K08AG029157-01; A15R01AG039384-03), American Federation for Aging Research, Lilly, Functional Neuro-modulation, and Pfizer, and has served as a consultant/advisor to Janssen, Pfizer, Lundbeck, and Abbvie.
Description of authors’ roles

M. Soto, B. Vellas, C. Lyketsos, and P. Rosenberg designed the manuscript. M. Soto wrote the manuscript. C. Lyketsos and P. Rosenberg assisted with writing the article. B. Vellas, C. Ballard, S. Andrieu, F. Nourhashemi, P. Robert, and PJ Ousset revised and corrected the manuscript.

Acknowledgment

This work was supported by NIH grant P50AG005146 to the Johns Hopkins Alzheimer’s Disease Research.

References


