ORIGINAL RESEARCH

*Naghmeh Foroutan, PhD^{1,2},
Robert B Hopkins, PhD^{1,2},
Jean-Eric Tarride, PhD^{1,2,3},
Ivan D Florez, MD, MSc^{1,4},
Mitchell Levine, MD, MSc^{1,2,3}*

¹ Department of Health Research Methods,
Evidence and Impact (HEI), McMaster
University, Hamilton, ON

² Programs for Assessment of Technology in
Health (PATH), St. Joseph Health Care
Hamilton, Hamilton, ON

³ Center for Health Economics and Policy
Analysis (CHEPA), McMaster University,
Hamilton, ON

⁴ Department of Pediatrics, University of
Antioquia, Medellín, Antioquia, Colombia

Safety and efficacy of active and passive immunotherapy in mild-to-moderate Alzheimer's disease: A systematic review and network meta-analysis

Abstract

Objective: The objective of this study was to systematically review and conduct a direct and network meta-analysis of randomized controlled trials that have examined the clinical safety and efficacy of using passive and active immunotherapies in Alzheimer's disease (AD).

Research questions:

- (1) Is amyloid-based immunotherapy in patients with mild-to-moderate AD associated with more efficacy benefits compared to placebo?
- (2) Which immunotherapy agent is associated with more comparative benefit?
- (3) Is passive or active immunotherapy associated with more benefits?

Data sources: A systematic review of published randomized controlled trials was performed in MEDLINE, EMBASE, PubMed and Cochrane library.

Review methods and meta-analysis: Two reviewers independently selected the studies, extracted the data and assessed risk of bias. Important AD cognitive scales as clinical efficacy outcomes were ADAS-cog, CDR and MMSE whereas edema, neoplasms and mortality were included as safety outcomes. A direct comparison meta-analysis using a random effect model and a network (direct and indirect) comparison was conducted to calculate mean differences in treatment effects, SUCRA and ranking probabilities for each medicine per safety and efficacy outcome. Quality of network results were assessed using GRADE methodology.

Principle findings: Thirteen RCT-assessed patients with mild-to-moderate AD were included in the final analysis. The results showed that immunotherapies compared with placebo produced a statistically, but not clinically significant, improvement in ADAS-cog (MD=-0.39; 95% CI -0.42, -0.35, P=0.00) and MMSE. In terms of safety, the rate of ARIA-E was significantly higher with monoclonal antibodies. Solanezumab and AN1792 (vaccine) were the drugs of choice both from efficacy and safety perspectives.

Conclusion: In terms of efficacy, the review showed a statistically, but not clinically significant, improvement in favor of immunotherapy versus placebo. Further clinical trials are required to demonstrate any cognitive benefits of immunotherapies in mild-to-moderate AD.

Correspondence to:

Naghmeh Foroutan
Email: foroutn@mcmaster.ca

*Manuscript submitted 31st July, 2018
Manuscript accepted 28th January, 2019*

Clin Invest Med 2019; 42 (1): E53-65.

There was an estimated 46.8 million people worldwide living with dementia in 2015. This number will almost double every 20 years, reaching 74.7 million in 2030 and 131.5 million in 2050 [1]. According to the National Institute of Aging, dementia is a brain disorder that affects communication and performance of daily activities and Alzheimer's disease (AD) is a form of dementia that specifically affects parts of the brain that control thought, memory and language. According to the Center for Disease Control, AD causes 50% to 70% of all dementia cases [2]. In 2011, 747,000 Canadians were living with AD and other dementias: that is 14.9% of Canadians 65 years of age and older. By 2031, it is expected that, this figure will increase to 1.4 million [3].

The growing incidence of AD has resulted in increasing demand to search for a medication that effectively targets the degenerative process. AD is associated with the accumulation of A β , an extracellular protein fragment that forms deposits around neurons. Those deposits, or plaques, trigger neuronal degeneration and attract inflammatory microglia, which further accelerate neuron death. Preventing the formation of A β plaques or clearing away A β before it reaches toxic levels is the central focus of AD therapeutic development. In principle, A β can be targeted with injected antibodies [4]. Two A β -binding antibodies, bapineuzumab (AAB-001) and solanezumab (LY2062430), have shown non-significant differences vs placebo in phase 3 testing to prevent AD progression in patients with mild to moderate disease [5,6].

An alternative approach consists of building up a patient's natural immune response against A β before the onset of disease. The simplest way to do this is with active immunotherapy or vaccination, in which A β is introduced to the peripheral immune system in a nontoxic form and stimulates a robust antibody-based response. This was the idea behind the AN1792 AD vaccine candidate. Vaccination with a full-length A β 42 peptide (AN1792) successfully elicited anti-A β antibodies in human subjects with AD, but was associated with meningoencephalitis, a neuroinflammatory condition thought to be caused by T cell activity in the brain, in about 6% of patients; thus, the phase 2 clinical trial was terminated before completion [7]. To avoid this safety issue, an aminoterminal A β 1-7 peptide conjugate, vanutide cridifcar (ACC-001), was designed and is currently in clinical development and recently passed phase 2 clinical trials [8,9]. Since then, academic researchers and companies have pursued a range of strategies to increase

antibody production and decrease inflammatory T cell activity in second-generation AD vaccine candidates [8,9].

The primary research question is "Is amyloid-based immunotherapy used in patients with the diagnosis of mild-to-moderate AD (status according to the National Institute of Neurologic and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association criteria) associated with more benefits than harms compared to placebo?" The secondary research questions were "Which immunotherapy agent is associated with more comparative benefit (intra-class ranking in terms of clinical efficacy)?" and "Is passive, or active immunotherapy associated with more benefits?"

Methods

Systematic search of the literature

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines [10]. This report adheres to the recommendations of the PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions [11].

The search strategy was performed by an information specialist (KC) in collaboration with the primary investigator (NF) in MEDLINE and EMBASE using the Ovid interface, and in the Cochrane Central Register of Controlled Trials (CENTRAL). Clinical trials were followed in clinicaltrials.gov database and the grey literature was searched for news or specific announcements about the medications under research and development. Search filters for the concepts of AD; immunotherapy and A β peptide were constructed using a combination of database MESH headings and text words. Database terms were adapted for each electronic database searched, no language restrictions were applied.

The research question and study eligibility criteria for this systematic review are based upon the following PICOS descriptions [12]:

Population: All adults with clinical diagnosis of mild-to-moderate AD according to standardized diagnostic criteria (according to the National Institute of Neurologic and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association).

Intervention: Any medicines used for active or passive immunotherapy (bapineuzumab, AAB-001, MABT5102A, ponezumab, PF-04360365, gantenerumab, RO4909832, solanezumab, LY2062430, BAN2401, GSK933776,

AFFITOPE, AN1792, CAD106, semagacestat, LY450139, γ -secretase inhibitor, tramiprosate, 3APS, homotaurine, ALZHEMED, IVIG)

Comparators: Immunotherapies—passive and active (inter-class comparison; head-to-head studies) or placebo

Outcomes:

- Alzheimer's disease Assessment Scale-cognitive subscales (ADAS-cog) [13]:
 - ADAS-cog14; range 0-90, with higher scores indicating greater cognitive impairment
 - ADAS-cog11; range 0-70, with higher scores indicating greater cognitive impairment;
- Clinical Dementia Rating–Sum of Boxes: range 0-18, with higher scores indicating worse functioning [13];
- Mini-Mental State examination (MMSE; score range 0-30, with higher scores indicating better cognitive function) mild AD (MMSE score of 20-26) and moderate AD (MMSE score of 16-19);
- amyloid related imaging abnormalities with edema (ARIA-E);
- neoplasms benign, malignant and unspecified; and
- mortality.

Study types: Published or unpublished studies that were completed (with results) from phase 2 and phase 3 randomized controlled trials (RCTs) on humans, in any language. Studies were excluded if they were (1) post hoc analyses of previous RCTs, (2) non-RCTs, (3) studies with same population (duplication) and (4) trials that did not measure primary clinical outcomes or did not report required data.

Study selection, data abstraction and quality assessment

Two authors (NF and RH) independently selected studies for inclusion in this review. From the titles and abstracts of all studies identified by the electronic search, those that clearly did not satisfy inclusion criteria were excluded. Full texts were retrieved when either one or both of the reviewers decided to include it. Full copies of the studies considered eligible were then obtained and reviewed independently by the two reviewers to identify those suitable for inclusion in the analyses. Data extraction of included studies was also performed independently by the same reviewers. Disagreements or uncertainty were resolved by consensus.

The two review authors also independently assessed the risk of bias for included RCTs using the Cochrane Collaboration's Risk of Bias tool and abstracted data using a

study-specific data extraction form. Again, disagreements were resolved by consensus.

Measures of treatment effect

The pairwise relative treatment effects of the competing interventions were estimated by calculating effect sizes as risk ratios (RRs) with 95% confidence intervals (CIs) for dichotomous outcome data (e.g., mortality), and by calculating mean differences with 95% CIs for continuous outcome data. For scales such as ADAS-cog, where a higher score indicates an increased disease severity, a negative mean difference vs placebo or other drugs is favorable and the drug will have a low rank.

Data synthesis, direct and network meta-analyses

A standard pairwise meta-analyses on the results was performed when data from included studies were available for any treatment comparison. A random-effects meta-analysis using an inverse variance weighting method was performed. Statistical heterogeneity was assessed by calculating separate heterogeneity variances for each pairwise comparison and by using Chi2 statistic and its P value, and the I2 statistic to quantify the percentage of variability that is due to true differences among studies rather than sampling error [14,15].

A random-effects network meta-analysis (NMA) was performed to obtain estimates for each comparison. In the absence of direct evidence for a given comparison, the indirect comparison provided the estimate. In the presence of direct evidence, the NMA is a combined estimate (i.e., direct and indirect evidence) [16]. In NMA, a common estimate was assumed for heterogeneity variance across different comparisons. We planned to combine direct and indirect estimates if the assumptions of between-study homogeneity, incoherence and transitivity across treatment comparisons were justifiable. Nonetheless, all the loops except one were open (all the interventions were only compared to placebo, none to each other), which meant that for most of the comparisons, there was no direct evidence. The only closed loop available was composed by three interventions that were only tested in a three-arms RCT [8], and since loop incoherence cannot occur within a multi-arm trial [17], we did not assess statistical incoherence.

A sensitivity analysis was performed to assess whether the findings of this review were robust to the decisions made in the process of obtaining them by conducting the following procedure: re-analysis excluding studies according

to study quality issues, re-analysis including those with low or high risk of bias, re-analysis without imputing data for the missing participants and re-analysis using a random-effect model.

The ranking probabilities for each treatment were estimated. This is the probability that each treatment is the best, the second best, the third best, etc. in the network, based on the magnitude of the effects. Based on these probabilities we calculated the surface under the cumulative ranking curve (SUCRA), which is a summary of the rankings [18]. Surface under the cumulative ranking curve were expressed as a percentage and interpreted as the percentage of efficacy or safety of a treatment that would be ranked first without uncertainty. All the analyses were performed with STATA version 14 (StataCorp 2015).

Rating the confidence in estimates of the effect

We assessed the confidence in the estimates (also called quality of the evidence or certainty on the evidence) for ADAS-cog and mortality outcomes using the GRADE approach [19]. The confidence in the estimates was based on four levels: high, moderate; low; and very low. For the assessment of the confidence on the direct comparisons estimates, we rated both outcomes based on the traditional GRADE categories: risk of bias; imprecision; inconsistency; and publication bias [20-24]. Furthermore, following the GRADE guidance [25], in a second step we removed the imprecision criterion to create direct GRADE quality assessments based only on four criteria (not rating down if imprecision was present, because this criterion was applied to the NMA estimate assessment, as is detailed below) to be used to inform the NMA estimates quality assessments.

We rated the confidence in the indirect comparisons focusing our assessments on loops connected to the interventions of interest through only one other intervention. A loop of evidence exists when two or more direct comparisons contribute to an indirect estimate [26]; for example, when there are at least one study comparing intervention A vs intervention B, and at least one study comparing intervention A vs intervention C, the information from these two comparisons can be used to calculate an indirect estimate (indirect evidence) of a comparison B vs C. A loop is considered closed if direct evidence exists between B vs C (at least one study comparing B vs C) and will be considered an open loop when this direct evidence does not exist.

For the confidence assessment we followed the guidance provided by Puhan et al. [27] and Brignardello et al. [25]. In the previous example, we had evidence A vs B (AB) and A vs C (AC), and we indirectly estimated the effects of B vs C (BC). The indirect comparison confidence was the lowest of the confidence ratings we assigned to the two contributing direct comparisons [27]; for instance, if AB had a confidence rated as moderate and BC had a confidence rated as high, the associated indirect comparison, AC, was judged as moderate confidence.

Furthermore, we had planned to rate down confidence in the indirect comparisons estimates further if we had a strong suspicion of intransitivity; however, transitivity was not suspected in any of the comparisons. Transitivity, also called similarity [28], is the assumption that an indirect comparison is a valid method to compare two treatments because the studies are sufficiently similar in important clinical and methodological characteristics; in other words, that they are similar in their distributions of effect modifiers [29,30]. For instance, following the scenario described above, in which we have evidence AB and BC but not AC, and the studies comparing AB are substantially different from studies comparing BC, in terms of characteristics of the population or the severity of the disease, we could define that there is intransitivity and, therefore, we would have reduced the confidence in the indirect estimate AC. As stated above, we did not identify intransitivity in any of the loops.

Lastly, we planned to rate the confidence in the NMA estimate for any pairwise comparison using the higher of the confidence rating amongst the contributing direct and indirect comparisons, and to rate down confidence in the network estimate when direct and indirect estimates have incoherence and/or imprecision [27]. Nevertheless, since there were no closed loops available, except one created by a three-arms trial, incoherence was not assessed in the comparisons and not considered in the closed loop.

Results

Search results and study selection

We retrieved references from Ovid MEDLINE (n= 509), Ovid EMBASE (n= 574), PubMed (n=31) and the CENTRAL (n=12) (Supplementary Material A). Moreover, six additional RCTs were identified from previous reviews reference list (Figure 1). As shown in Figure 1 (PRISMA flowchart), out of 998 references that were assessed in level one screening (title and abstract), 43 fulfilled the first level

inclusion criteria (Supplementary Material B.a) and assessed in level 2 (full-text) screening phase (Supplementary Material B.b). Ten articles (13 RCTs) [5-8,31-36] fulfilled all the inclusion criteria and none of the exclusion criteria. The reasons for excluding 33 potential RCTs were as follows: post hoc analyses of previous RCTs; non-RCTs; studies with same population; and trials that did not measure primary clinical outcomes or did not report required data.

Study and patient characteristics

Table 1 summarizes the characteristics of the ten articles including thirteen phase 2 and phase 3 RCTs (three articles included two studies each). For drug efficacy assessment, 6,699, 6,559 and 5,802 participants were included for ADAS-cog, CDR-SB and MMSE outcome measure scales, respectively. Follow-up duration ranged from 3 to 20 months and study size ranged from 27 to 1,331 subjects.

FIGURE 1. Article selection flowchart

TABLE 1. Characteristics of patients and studies

Author, year, country	Patient characteristics*	Intervention	Comparator	Phase	Duration (months)	Number of randomized patients	Drug	Placebo
Aisen, 2011, North America (USA, Canada)	Aged ≥ 50 years mild-to-moderate AD, Probable AD; MMSE scores of 16 to 26	tramiprosate (3APS, homotaurine, ALZHEMED) oral tab: 150 mg	Placebo	3	20	1,052	229	248
Dodel, 2013, USA, Germany	Aged ≥ 50 years mild-to-moderate AD, Probable AD; MMSE scores of 16 to 26	IVIg 0.4 g/kg	Placebo	2	5.5	55	6	5
Doody, 2013, USA	Aged ≥ 50 years mild-to-moderate AD, Probable AD; MMSE scores of 16 to 26	semagacestat 140 mg	Placebo	3	19	1,030	497	486
Doody, 2014, USA	Aged ≥ 50 years mild-to-moderate AD, Probable AD; MMSE scores of 15 to 26	solanezumab 400 mg IV	Placebo	Two phase 3	18	1,012 (Expedition1) 1,040 (Expedition2)	506 (Expedition1) 521 (Expedition2)	506 (Expedition1) 519 (Expedition2)
Farlow, 2012, USA	Aged ≥ 50 years mild-to-moderate AD, Probable AD; MMSE scores of 15 to 26	solanezumab	Placebo	2	4	52	10	10
Fleisher, 2008, USA	Aged ≥ 50 years mild-to-moderate AD, Probable AD; MMSE scores of 15 to 26	semagacestat 140 mg	Placebo	2	3.5	51	13	12
Gilman, 2005, USA, EU	Aged ≥ 50 years mild-to-moderate AD, Probable AD; MMSE scores of 15 to 26	AN1792; 225 mcg and QS-21 50 mcg	Placebo	2	15	372	44	53
Salloway, 2009, USA	Aged ≥ 50 years mild-to-moderate AD, Probable AD; MMSE scores of 16 to 26	bapineuzumab 0.5, 1.0 mg/kg	Placebo	2		234	26	22
Salloway, 2014, USA, Canada, Germany, Austria	Aged ≥ 50 years mild-to-moderate AD, Probable AD; MMSE scores of 15 to 26	bapineuzumab 0.5, 1.0 mg/kg	Placebo	Two phase 3	19.5	APOE 4 carriers: 1,121 APOE 4 Non-carriers: 1331	314 314	493 493

Pasquier, 2016, USA, European Union	Aged 50-85 years, mild-to-moderate AD, Probable AD; MMSE scores of 16 to 26	ACC-001 +/- QS-21	Placebo	Two phase 2a	12	245	184	61
--	---	-------------------	---------	--------------	----	-----	-----	----

*Status according to the National Institute of Neurologic and Communicative Disorders and Stroke–Alzheimer’s Disease and Related Disorders Association criteria.

Abbreviation: MMSE, Mini-Mental State Examination score range, 0 to 30, with higher scores indicating better cognitive function mild Alzheimer’s disease (MMSE score of 20 to 26) and moderate Alzheimer’s disease (MMSE score of 16 to 19)

TABLE 2. Results of quality assessment of thirteen randomized controlled trials

Author, year	Sequence generation	Allocation concealment	Blinding	Blinding outcome assessor	Percentage that completed the trial (incomplete outcome data)*	Selective outcome reporting	Other bias
Aisen, 2011	Randomization list issued by an independent biostatistician, using a computer random number generator, and balanced to ensure	Yes	Double-blind	Unclear	75 (Low)	Low	Low
Dodel, 2013	Computer-generated randomization list created by the contract research organization with SAS (version 9.1.3). Patients were allocated through an interactive web response service in block sizes of eight.	Yes	Double-blind	Blinded	62 (High)	Low	Low
Doody, 2013	Unclear	Unclear	Double-blind	Unclear	30 (High)	Low	Low
Doody, 2014	Unclear	Unclear	Double-blind	Unclear	73 (Low)	Low	Low
Farlow, 2012	Interactive voice response system - vendor generated the treatment assignments	Yes	Double-blind	Unclear	96 (Low)	Low	Low
Fleisher, 2008	Telephone-based interactive voice response system	Yes	Double-blind	Unclear	81 (Low)	Low	Low
Gilman, 2005	Computerized, random-number generator	Yes	Double-blind,	Unclear	80 (Low)	Low	Low
Salloway, 2009	Adaptive stratified randomization	Unclear	Double-blind	Unclear	68 (Low)	Low	Low

Salloway, 2014	Randomization was stratified according to the use or nonuse of a cholinesterase inhibitor or memantine, baseline MMSE total score (16 to 21 vs 22 to 26), participation in a substudy, and, in the carrier study, APOE ε4 copy number (1 vs 2).	Unclear	Double-blind (adaptive design)	Unclear	97 (Low)	Low	suspected for type 1 systematic error
Pasquier, 2016, USA, European Union	Patients were randomized to receive either placebo or one of three dose levels of ACC-001 (3, 10, or 30 mcg) with or without QS-21 adjuvant (50 mcg) in an escalating cohort design.	Unclear	Double-blind†	Blinded	79 (Low)	Low	Low

*High, high risk of bias; low, low risk of bias

†All parties were blinded to treatment allocation except for the dispensing pharmacists, who were not involved in patient evaluation.

Methodological quality and risk of bias

As observed in Table 2, the studies were found to be of “moderate” to “high risk of bias”. Attrition bias (losses of follow-up) was >20% in two studies [32,33]. Detection bias (blinding of assessors, which is important in the case of ADAS-cog, but not mortality) was also observed in most of the studies.

Efficacy assessment

Alzheimer's disease Assessment Scale-cognitive subscales (ADAS-cog):

In terms of direct evidence, as per the efficacy primary ADAS-cog outcome, nine articles [5-8,31,32,34-36] evaluated the efficacy of AN1792 (N= 44), bapineuzumab 0.5 mg/kg (N= 990), bapineuzumab 1 mg/kg (N= 340), solanezumab 400 mg (N= 1037), semagacestat 140 mg (N= 510), IVIG 0.4 g/kg (N= 6), 3APS 150 mg (N= 229) and ACC-001 +/- QS-21 (N= 184) vs placebo. The meta-analysis results showed a statistically significant change from baseline regarding ADAS-cog values (mean difference=-0.39; 95% CI -0.42, -0.35, P=0.001) in favor of immunotherapies; however, the ADAS-Cog is a detailed cognitive assessment for dementia for which a four-point difference between treatment groups is required to be considered a significant difference in the clinical practice setting which was not achieved in the present analysis (Figure 2).

Figure 3 shows the results of network plot analysis making comparisons between immunotherapies and placebo. The results of NMA (interval plot) are shown in figure B.1 (Supplementary Material B). Consequently, the results of “ranking probabilities for competing treatments” or SUCRA command, solanezumab 400 mg is reported as the best therapeutic choice having measured cognitive outcomes of patients with the ADAS-cog scale (Figure B.2, Supplementary Material B). There was no change in ranking results of low attrition bias studies after the sensitivity analysis.

Clinical Dementia Rating–Sum of Boxes (CDR-SB): The direct meta-analysis results showed a statistically but not clinically significant difference in favor of placebo vs immunotherapies (mean difference=0.11, 95% CI 0.10, 0.12, P=0.001) from the six studies (eleven comparisons) evaluating efficacy of AN1792 (N= 44), bapineuzumab 0.5 mg/kg (N= 972), bapineuzumab 1 mg/kg (N= 307), solanezumab 400 mg (N= 1,027), semagacestat 140 mg (N= 494), IVIG 0.4 g/kg (N= 6), 3APS 150 mg (N= 233) and ACC-001 +/- QS-21 (N= 184) in patients with the CDR-SB scale. No change was observed in the results following a sensitivity analysis assessing studies with low vs high risk of bias. The results of “ranking probabilities for competing treatments” or SUCRA command, an active immunotherapy

FIGURE 2. Meta-analysis (forest plot) results for the ADAS-cog scale

(AN1792 225 mcg) and solanezumab 400 mg as a passive immunotherapy have been reported as the best alternatives (Figure B.3, Supplementary Material B). As with the ADAS-cog, the higher the score; the worst the cognitive condition.

Mini-Mental State examination (MMSE): Six articles [5-8,32,33] reported efficacy results for AN1792 (N= 42), bapineuzumab 0.5 mg/kg (N= 972), bapineuzumab 1 mg/kg (N= 307), solanezumab 400 mg (N= 1,027), semagacestat 140 mg (N= 303), IVIG 0.4 g/kg (N= 6) and ACC-001 +/- QS-21 (N= 155) using the MMSE scale. As illustrated in

FIGURE 3. Network meta-analysis plot for the ADAS-cog scale

Abbreviations: ACC-001, vanutide cridificar; BAPZ, bapineuzumab; IVIG, intravenous immunoglobulin; Semagt, semagacestat; Solz, solanezumab; 3APS, tramiprosate (homotaurine, ALZHEMED)

Figure B.4 (Supplementary Material B), the direct meta-analysis results showed a statistically significant difference between drug therapies and placebo (MD=0.04, 95% CI 0.02, 0.05, $P=0.00$). According to the results for the SUCRA command, solanezumab 400 mg has been reported as the best therapeutic choice using the MMSE scale (Figure B.4, Supplementary Material B). No change in ranking was observed in the therapeutic ranking results following a sensitivity analysis assessing studies with low vs high loss of follow-up [32,33].

Safety assessment

Amyloid related imaging abnormalities with edema (ARIA-E): Three studies [5,6] reported significantly higher rate of ARIA-E as RR= 9.3 (95% CI, 3.56, 24.35; $P<.01$), vs placebo in 4856 patients using monoclonal antibodies (bapineuzumab 0.5, N=995, 28%; bapineuzumab 1 mg/kg, N= 329, 13%; solanezumab 400 mg, N= 1,027, 1%) in mild-to-moderate AD. Network comparison results (SUCRA)

showed that with respect to ARIA-E, solanezumab 400 mg is a safer alternative compare to bapineuzumab 0.5 and 1 mg/kg in mild-to-moderate AD patients.

Neoplasms-benign, malignant and unspecified:

The direct meta-analysis results showed a non-significant difference in rate of neoplasms between drug therapies and placebo (RR=1.14, 95% CI 0.53, 2.5, $P=0.74$) from four studies [5,6,33] reporting rate of neoplasms in bapineuzumab 0.5 mg/kg (N= 785, 4.7%), 1 mg/kg (N= 329, 1.8%), solanezumab 400 mg (N= 1,027, 3.8%), semagacestat 140 mg (N= 527, 3%), vs placebo in 5,915 patients with mild-to-moderate AD. For neoplasms as a severe adverse event, the results for the SUCRA command reported bapineuzumab (0.5 and 1 mg/kg) as safer therapeutic alternatives compared with solanezumab 400 mg and semagacestat 140 mg.

Mortality:

Six articles [5-8,31,33] reported death as a drug related severe adverse event: AN1792 225 mcg (N=300, 1.7%); bapineuzumab 0.5 mg/kg (N= 995, 3.5%); bapineuzumab 1 mg/kg (N= 329, 2%); solanezumab 400 mg (N= 1,027, 2.3%); semagacestat 140 mg (N= 527, 2.7%); IVIG 0.4 g/kg (N= 42, 0%); 3APS 150 mg (N= 347, 0.3%); and ACC-001 +/- QS-21 (N= 154, 1.3%). There was a non-significant difference between the rate of death in the immunotherapy group and the placebo (RR=1.40, 95% CI 0.96, 2.02, $P=0.07$) with the heterogeneity zero ($I^2=0.0\%$). The results of NMA and ranking probabilities of treatments, AN1792 225 mcg as an active immunotherapy resulted in the lowest rate of death and the second safer alternative as passive immunotherapy is solanezumab 400 mg (Figure B.5, Supplementary Material B).

GRADE assessment

Overall the quality of the evidence was low to very low in most of comparisons for ADAS-cog changes and mortality. Tables C.1 (Supplementary Material C) represent a summary

of the direct and NMA effect estimates and 95%CI for each comparison in terms of ADAS-cog mean difference, and the quality of the evidence assessment. Only the comparison IVIG 0.4 g vs placebo was rated as moderate quality, while semagacestat 140 mg vs placebo, ACC001+QS vs placebo, ACC001 vs placebo, and ACC001+QS vs ACC001 were rated as low quality and the rest of comparisons were rated as of very low quality. Table C.2 (Supplementary Material C) summarizes the mortality outcomes in terms of RR and 95% CI. Only in the comparisons of ANT192 vs placebo, 3APS150 vs placebo, ACC001+QS vs placebo, ACC001 vs placebo and ACC001+QS vs ACC001 the evidence was rated as moderate, while the rest of comparisons were low or very low quality.

Discussion

To assess the potential benefits, safety and efficacy of immunotherapies reported in phase 2 and phase 3 clinical trials over last decade, a systematic review and network meta-analysis was performed to provide clinicians and researchers with best evidence and assist in determining future research and designing clinical trials.

Although the network meta-analysis was not ideal given the lack of head-to-head studies among the interventions (i.e., only against placebo; a star-shaped network), except for one three-arm trial, interesting results were found. In efficacy assessment, the direct meta-analyses showed that immunotherapy vs placebo was statistically but not clinically significant in change from baseline in ADAS-cog and MMSE results [32,33]. No clinical benefit in favor of immunotherapy was observed with the CDR-SB scale. The best alternative as a result of ranking probabilities was solanezumab 400 mg every four weeks (Q 4 W) for improving ADAS-cog, CDR-SB and MMSE scales. The use of AN1792 225 mcg as a vaccine, which is still under development due to serious adverse events (meningoencephalitis), would be the second ranked active immunotherapy alternative for improving CDR-SB.

Three clinically-important serious adverse events were selected to be compared among immunotherapies and placebo (inter- and intra-group comparison). With respect to neoplasms and mortality, no significant differences were observed in direct comparisons vs placebo. Although the change in the rate of mortality outcome was not statistically significant, the study results showed a trend towards increasing mortality, and that significance might not be achieved because of a low sample size. Perhaps further studies with

more events, would have confirmed this trend. There was a significant higher rate of amyloid-related imaging abnormalities with edema (ARIA-E) observed in immunotherapy group vs placebo (RR= 9.3; 95%CI, 3.56, 24.35).

According to the results of ranking probability matrix, the best alternatives were as follows: bapineuzumab 0.5 mg/kg, bapineuzumab 1 mg/kg, solanezumab 400 mg and AN1792 225 mcg, which showed the lowest rates of neoplasms, ARIA-E and mortality. Considering that AN1792 225 mcg had a specific serious adverse event (meningoencephalitis), a safer alternative was solanezumab 400 mg, which showed the second lowest mortality.

Rankings and effect estimates should be analyzed together with the quality of the evidence. Several limitations in terms of risk of bias were found among most of the studies (lack of blinding of outcome assessors, significant losses of follow-up and allocation concealment flaws). This, along with additional limitations related to imprecision, meant that the quality of evidence for changes in ADAS-cog and the effect on mortality were judged to be very low for almost all the indirect comparisons obtained. For the direct comparisons of interventions vs placebo, only one comparison showed moderate quality in the ADAS-cog outcome, and four comparisons in the mortality outcome. Although the rankings provide us an idea of the effectiveness, the uncertainty around the estimates was quite high given the quality of the evidence.

The efficacy of IVIG 0.4 g/kg to reduce the ADAS-cog, was the only effect rated with moderate quality; therefore, the uncertainty around this was the lowest, but the results of this intervention was not different enough from the placebo to produce changes in the scores. Results with very low quality evidence were largely explained by the high risk of bias analyses.

The quality of evidence assessed by GRADE considered several criteria for the rating of evidence, and the presence of only indirect evidence (except for ACC001+QS vs ACC001) reduced the quality of evidence and resulted in the inability to calculate a NMA estimate. This was another reason for the final very low quality among almost all the indirect comparisons available.

The most important limitations observed in this review were 1) the absence of head-to-head trials, leading to mostly indirect estimates for all the comparisons among the interventions, 2) the short-term period of follow-up in the studies, 3) the losses of follow-up in some RCTs leading to a

high risk of bias, 4) the absence of phase 3 clinical trials for some medications, 5) the termination of one RCT before completion due to severe adverse events, 6) the high rates of heterogeneity, mainly due to different follow-up periods and 7) the low number of included studies for each medication.

Summary

This analysis provides researchers with a clearer idea of the safety and efficacy of immunotherapies for AD that is based on best available evidence. Previously, there was no systematic review or meta-analysis making inter- and intra-group comparisons of safety and efficacy of amyloid-based immunotherapies in AD. We have systematically reviewed all the available published RCTs that examined the harm and benefit of using immunotherapy for AD. We summarized all the evidence and determined the comparative effectiveness of all the interventions. The present study is the first systematic review network meta-analysis of efficacy and safety of immunotherapies in the treatment of patients with mild-to-moderate AD. We limited the inclusion criteria to only published phase 2 and phase 3 RCTs to obtain the best evidence for addressing this important clinical issue in AD management.

Conclusions

This review results showed no clinically significant improvement in any AD-specific outcome measure scales (ADAS-cog, CDR-SB and MMSE) in favor of immunotherapies (passive and active) in mild-to-moderate AD. Solanezumab 400 mg was the drug of choice from both efficacy (ADAS-cog and MMSE) and safety (ARIA-E) perspectives and bapineuzumab 1 mg/kg showed the lowest risk of neoplasms compare with alternatives. Active immunotherapy (AN1792 225 mcg) showed better outcomes in both efficacy (CDR-SB) and mortality (vs passive immunotherapy) but had a potential specific serious adverse event (meningoencephalitis). To address this safety issue, vanutide cridifcar (ACC-001) has been developed and has recently passed phase 2 clinical trials, showing an acceptable safety profile. Further high quality phase three RCTs, especially for vaccines and IVIG, are required to detect any meaningful differences in the cognitive endpoints using immunotherapies in mild-to-moderate AD.

Acknowledgments

Authors would like to say thank you to Kaitryn Campbell for her contribution and assistance in this publication.

Appendix

Supplementary Material A, B and C—available from the corresponding author upon request.

References

1. Campbell, D.M., et al., Increasing the use of evidence in health policy: practice and views of policy makers and researchers. Aust New Zealand Health Policy, 2009. 6: 21.
2. What Is Alzheimer's? [cited 2019 Jan 4] www.alz.org/alzheimers-dementia/what-is-alzheimers.
3. Alzheimer's disease and other dementias in Canada, in Statistics Canada. 2016.
4. Relkin, N., Clinical trials of intravenous immunoglobulin for Alzheimer's disease. J Clin Immunol, 2014. 34 Suppl 1: S74-9.
5. Doody, R.S., et al., Phase 3 Trials of Solanezumab for Mild-to-Moderate Alzheimer's Disease. N Engl J Med, 2014. 370(4): 311-321.
6. Salloway, S., et al., Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. N Engl J Med, 2014. 370(4): 322-33.
7. Gilman, S., et al., Clinical effects of Abeta immunization (AN1792) in patients with AD in an interrupted trial. Neurology, 2005. 64(9): 1553-62.
8. Pasquier, F., et al., Two Phase 2 Multiple Ascending-Dose Studies of Vanutide Crdifcar (ACC-001) and QS-21 Adjuvant in Mild-to-Moderate Alzheimer's Disease. J Alzheimers Dis, 2016. 51(4): 1131-43.
9. Arai, H., H. Suzuki, and T. Yoshiyama, Vanutide crdifcar and the QS-21 Adjuvant in Japanese subjects with mild to moderate Alzheimer's disease: Results from two phase 2 studies. Current Alzheimer Research, 2015. 12(3): 242-254.
10. Moher, D., et al., Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg. 8(5): 336-41.
11. Hutton, B., et al., The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. Annals of Internal Medicine, 2015. 162(11): 777-784.
12. Guyatt, G.H., et al., GRADE guidelines: 2. Framing the question and deciding on important outcomes. J Clin Epidemiol. 64(4): p. 395-400.

13. Sheehan, B., Assessment scales in dementia. *Ther Adv Neurol Disord*, 2012. 5(6): 349-58.
14. Higgins, J.P., et al., Measuring inconsistency in meta-analyses. *Bmj*, 2003. 327(7414): 557-60.
15. Higgins, J.P.T. and S.G. Thompson, Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 2002. 21(11): 1539-58.
16. Mills, E.J., K. Thorlund, and J.P. Ioannidis, Demystifying trial networks and network meta-analysis. *Bmj*, 2013. 346: f2914.
17. Higgins, J.P., et al., Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Res Synth Methods*, 2012. 3(2): 98-110.
18. Chaimani A, H.J., Mavridis D, Spyridonos P, Salanti G, Graphical Tools for Network Meta-Analysis in STATA. *PLoS ONE* 2013. 8(10): p. e76654.
19. Guyatt, G., et al., GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. *Journal of clinical epidemiology*, 2011. 64(4): 383-94.
20. Balshem, H., et al., GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol*, 2011. 64(4): 401-6.
21. Guyatt, G.H., et al., GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). *J Clin Epidemiol*. 2011. 64(4): 407-15.
22. Guyatt, G.H., et al., GRADE guidelines 6. Rating the quality of evidence—imprecision. *J Clin Epidemiol*. 2011. 64(12): 1283-93.
23. Guyatt, G.H., et al., GRADE guidelines: 7. Rating the quality of evidence—inconsistency. *J Clin Epidemiol*. 2011. 64(12): 1294-302.
24. Guyatt, G.H., et al., GRADE guidelines: 5. Rating the quality of evidence—publication bias. *J Clin Epidemiol*. 2011. 64(12): 1277-82.
25. Brignardello-Petersen, R., et al., Advances in the GRADE approach to rate the certainty in estimates from a network meta-analysis. *J Clin Epidemiol*, 2018. 93: 36-44.
26. Al Khalifah, R., et al., Network meta-analysis: users' guide for pediatricians. *BMC Pediatr*, 2018. 18(1): 180.
27. Puhan, M.A., et al., A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. *BMJ*, 2014. 349: g5630.
28. Donegan, S., et al., Indirect comparisons: a review of reporting and methodological quality. *PLoS One*, 2010. 5(11): e11054.
29. Baker, S.G. and B.S. Kramer, The transitive fallacy for randomized trials: if A bests B and B bests C in separate trials, is A better than C? *BMC Med Res Methodol*, 2002. 2: 13.
30. Cipriani, A., et al., Conceptual and technical challenges in network meta-analysis. *Ann Intern Med*, 2013. 159(2): 130-7.
31. Aisen, P.S., et al., Tramiprosate in mild-to-moderate Alzheimer's disease - a randomized, double-blind, placebo-controlled, multi-centre study (the Alphase Study). *Arch Med Sci*, 2011. 7(1): 102-11.
32. Dodel, R., et al., Intravenous immunoglobulin for treatment of mild-to-moderate Alzheimer's disease: a phase 2, randomised, double-blind, placebo-controlled, dose-finding trial. *Lancet Neurol*, 2013. 12(3): 233-43.
33. Doody, R.S., et al., A phase 3 trial of semagacestat for treatment of Alzheimer's disease. *N Engl J Med*, 2013. 369(4): 341-50.
34. Farlow, M., et al., Safety and biomarker effects of solanezumab in patients with Alzheimer's disease. *Alzheimer's & Dementia*, 2012. 8(4): 261-71.
35. Fleisher, A.S., et al., Phase 2 safety trial targeting amyloid beta production with a gamma-secretase inhibitor in Alzheimer disease. *Arch Neurol*, 2008. 65(8): 1031-8.
36. Salloway, S., et al., A phase 2 multiple ascending dose trial of bapineuzumab in mild to moderate Alzheimer disease. *Neurology*, 2009. 73(24): 2061-70.