

- 1 Number of words in the Abstract: 150
- 2 Number of words in the Introduction: 508
- 3 Number of words in the Discussion: 1156
- 4 Number of figures: 1
- 5 Number of tables: 3
- 6

7 **Donanemab outperformed Aducanumab and Lecanemab on cognitive, but not on**
8 **biomarker and safety outcomes: systematic review, frequentist and Bayesian**
9 **network meta-analyses**

10 Danko Jeremic, PhD, Juan D. Navarro-López[†], PhD, Lydia Jiménez-Díaz[†], PhD

11 [†]These authors share last senior and corresponding authorship.

12

13 NeuroPhysiology & Behavior Lab, Institute of Biomedicine (IB-UCLM), School of

14 Medicine of Ciudad Real, 13005, University of Castilla-La Mancha, Spain.

15

16 *Correspondence to:

17 J.D. Navarro-Lopez and L. Jimenez-Diaz

18 Neurophysiology and Behavior Lab

19 Universidad Castilla-La Mancha

20 13071-Ciudad Real

21 Phone: + 34-926295300 ext 6838

22 Fax: + 34-926295300

23 E-mail: Juan.Navarro@uclm.es; Lydia.Jimenez@uclm.es

24

25 **Keywords:** Network Meta-Analysis; Bayesian; Frequentist; Alzheimer's; Anti-amyloid;

26 Aducanumab; Lecanemab; Donanemab;

27 **ABSTRACT**

28

29 **INTRODUCTION**

30 Questions remain regarding safety and clinical relevance of anti-amyloid antibodies in
31 Alzheimer's disease (AD), with no scientific basis for choosing between different
32 therapies.

33 **METHODS**

34 Systematic review, frequentist and Bayesian network meta-analyses of phase III
35 randomized placebo-controlled trials were performed to comparatively evaluate
36 cognitive, functional and biomarker efficacy and safety of anti-amyloid antibodies in
37 sporadic AD. Treatments were ranked with P- and SUCRA scores, with rank robustness
38 measured by Cohen's *kappa*, and uncertainty in ranking probabilities estimated with
39 Shannon's normalized entropy.

40 **RESULTS**

41 Based on data from 16,971 patients (16 studies), we found Donanemab the best-ranked
42 antibody on cognitive measures. Lecanemab was the most effective at reducing amyloid
43 burden. Caution is needed concerning brain edema and microbleeding, with clinically
44 important risks for Donanemab, Aducanumab and Lecanemab.

45 **DISCUSSION**

46 Risk/benefit profile of anti-amyloid antibodies remains unfavorable. Patients in
47 Donanemab study were stratified by *tau* load, with greater effects observed in
48 low/medium *tau* population.

49

50 **Highlights**

- 51 • No single therapy ranked the best among all outcomes.
- 52 • Donanemab was the most effective antibody at reducing cognitive decline across
53 all primary outcomes, while Lecanemab ranked the highest on amyloid PET
54 removal.
- 55 • Consistently greater cognitive, functional and biomarker effects of Donanemab
56 were observed in patients with low/medium *tau* load, suggesting more promising
57 effects in earlier AD stages.
- 58 • All antibodies, except Solanezumab, were significantly less tolerable than
59 Placebo.
- 60 • The risk of cerebral edema and microbleeding may outweigh the benefits,
61 independently of APOE status.

62

63 1. BACKGROUND

64 Although known for a little more than a century, Alzheimer's disease (AD) became a
65 real-world problem in the last few decades as the most common form of dementia and
66 significant burden for our societies [1]. Until recently, symptomatic treatments were the
67 only therapies available for AD patients, and that situation remains in the majority of
68 countries in the world. The recent approvals of Aducanumab (Aduhelm[®], U.S) and
69 Lecanemab (Leqembi[®], U.S and Japan) have initiated the new era in AD therapeutics,
70 as these monoclonal antibodies have shown some potential to modify disease
71 progression by binding to amyloid- β (A β) peptide, toxic peptides thought to be essential
72 in the AD pathophysiology [2]. Both Aducanumab (epitope: 3–7) and Lecanemab
73 (epitope: 1–16) have high selectivity for N-terminus of full-length A β found in soluble
74 oligomers and fibrils, however, Lecanemab showed tenfold stronger binding to small
75 and large protofibrils than to fibrils [3, 4].

76 The regulatory approvals of Aduhelm[®] and Leqembi[®] were both promising and
77 controversial, and they came after more than a decade of investigating anti-A β vaccines
78 and antibodies in AD patients [5, 6]. Many other antibodies, such as Solanezumab
79 (specific to A β monomers) and Bapineuzumab (non-specific) were discontinued since
80 they failed to show beneficial effects in clinical trials. Another anti-A β antibody,
81 Donanemab, have gained significant attention recently and it is currently under review
82 for standard approval in the EU and the U.S. Instead of binding the N-terminus of full-
83 length A β , Donanemab targets the N-truncated pyroglutamate A β (at position 3), present
84 only in A β plaques. It is reported to rapidly induce plaque removal and reduce cognitive
85 decline in patients with early symptomatic AD [7, 8].

86 Despite clear advantages of disease-modifying therapies, substantial doubt remains
87 about both the clinical relevance and safety of anti-amyloid antibodies, and they still

88 have to demonstrate clinically meaningful effect in real life [9, 10]. We have previously
89 conducted the systematic review and web application meta-analysis of anti-A β
90 antibodies in phase III randomized placebo-controlled clinical trials (RCTs), showing
91 that Aducanumab and Lecanemab produced the most promising biomarker results
92 (reducing both A β and *tau* load) and improved cognitive outcomes by small effect size
93 when compared to control group. However, these cognitive effects were achieved at the
94 great expense of increasing adverse events, most importantly amyloid-related imaging
95 abnormalities (ARIA) in the form of cerebral edema (ARIA-E), hemorrhages and
96 localized superficial siderosis (ARIA-H) [11]. Our study had some limitations, as we
97 performed only the conventional (frequentist) meta-analysis method with DerSimonian
98 and Laird random-effect model. This did not allow us to estimate the relative effects and
99 compare the performance of these treatments. Also, we did not analyze amyloid PET
100 data from original studies, and we failed to include the effects of Donanemab, as phase
101 III trial results on this drug were not available at the time. Therefore, our aim in this
102 study was to update and expand our previous study by conducting random-effect
103 frequentist and Bayesian network meta-analyses (NMAs) of phase III RCTs in order to
104 compare the effects of anti-A β antibodies and rank their performance across cognitive,
105 functional, biomarker and safety outcomes.

106

107 2. METHODS

108 2.1. Search strategy

109 We followed the Preferred Reporting Items for Systematic Reviews and Meta-analyses
110 (PRISMA) – NMA guideline of health care interventions [12]. The comprehensive
111 search for published phase III trials was done on March 11, 2024, using the following
112 keywords: “Alzheimer’s”, “sporadic”, “mild cognitive impairment”, “phase 3”,
113 “monoclonal antibody”, “passive immunotherapy”, “Aducanumab”, “BIIB037”,
114 “Lecanemab”, “BAN-2401”, “Solanezumab”, “LY2062430”, “Bapineuzumab” and
115 “AAB-001”, “Donanemab” and “LY3002813”. The search was performed on Google
116 Scholar, PubMed, and clinical trial databases, including International Clinical Trials
117 Registry Platform (ICTRP) of the World Health Organization, the U.S. Clinical trial
118 Registry ClinicalTrials.gov, The EU Clinical Trial Registry and Australian New Zealand
119 Clinical Trials Registry (ANZCTR).

120 2.2. Quality assessment

121 Modified Jadad scale was used to assess the methodological quality of clinical trial
122 reports. The Jadad scale, also known as the Oxford quality scoring system has been
123 tested for reliability in different settings [13, 14] and includes the evaluation of study
124 randomization, blindness of patients and investigators, blindness in outcome
125 assessment, report of withdrawals, and dropouts [15]. In addition, we evaluated
126 reporting of the methodology used to assess adverse events (most notably ARIA),
127 whether ARIA is reported depending on the APOE genotype, and whether the methods
128 of the statistical analysis were described in the study. The studies were excluded if they
129 (1) had modified Jadad score < 3; (2) were not phase III RCTs; (3) did not include

130 patients with mild cognitive impairment (MCI) or sporadic AD; (4) tested less than 50
131 participants for cognitive and safety measures.

132 2.3. Risk of bias

133 The potential bias was classified as “low risk”, “with some concerns”, or “high risk, by
134 using Revised Cochrane risk-of-bias tool (RoB 2). The domains encompassed in RoB 2
135 analysis cover all types of bias that are currently understood to affect the results of
136 RCTs: (1) bias arising from the randomization process; (2) bias due to deviations from
137 intended interventions; (3) bias due to missing outcome data; (4) bias in measurement of
138 the outcome; and (5) bias in selection of the reported result [16]. Risk-of-bias
139 assessments were visualized using RoB 2 web app [17]. Finally, publication bias was
140 statistically evaluated for efficacy outcomes with more than ten studies [18] by using
141 Egger’s regression test [19], funnel plots, and the Duval & Tweedie’s Trim-and-Fill
142 procedure [20].

143 2.4. Grading the evidence and the impact of COVID-19 pandemic

144 To evaluate the quality of underlying evidence and the strength of recommendations in
145 this study, we followed Grades of Recommendation, Assessment, Development, and
146 Evaluation (GRADE) approach that focuses on six items: within-study bias, reporting
147 bias, indirectness, imprecision, heterogeneity, and incoherence. In addition, the impact
148 of COVID-19 pandemic on each study was evaluated by summarizing the number (%)
149 of patients affected by COVID-19 disease (Drug vs. Placebo) and how the pandemic
150 affected each stage of the trial, including adverse events and dropouts, missing doses,
151 delays and staffing difficulties, and whether the impact of COVID-19 was analyzed in
152 the main and sensitivity analyses.

153 2.5. Outcomes

154 Primary outcomes included efficacy measures reported by all or the majority of the
155 original studies. These were mean changes from baseline in cognitive function
156 evaluated by AD Assessment Scale-Cognitive Subscale (ADAS-Cog), Clinical
157 Dementia Rating scale-Sum of Boxes (CDR-SB), and Mini Mental State Examination
158 (MMSE).

159 Secondary outcomes were biomarker and safety measures reported by all or the majority
160 of studies. Biomarker outcomes included changes from baseline in amyloid burden on
161 PET (centiloid scale), cerebrospinal fluid (CSF) levels of $A\beta_{1-42}$ and p-tau (Thr₁₈₁), and
162 plasma *p-tau*. Safety measures were Serious adverse events, tolerability (treatment
163 discontinuations due to adverse events), total events of ARIA-E (cerebral edema),
164 cerebral microhemorrhages, total events of ARIA-H (including cerebral
165 microhemorrhages, macrohemorrhages and superficial siderosis), headaches, fatigue,
166 dizziness, syncope, nausea, falls, cardiac disorders, back pain, arthralgia,
167 nasopharyngitis, upper respiratory and urinary infections.

168 Tertiary outcomes included efficacy and safety measures reported by few studies only:
169 functional scale AD Cooperative Study-Activities of Daily Living in MCI (ADCS-
170 ADL-MCI), superficial siderosis (ARIA-H), infusion-related reactions and total events
171 of ARIA-E and ARIA-H in APOE- $\epsilon 4$ non-carriers and carriers (total and depending on
172 the $\epsilon 4$ allele dose).

173 The results of NMAs were summarized by League Tables with Standardized Mean
174 Differences (SMDs) calculated for efficacy (continuous) outcomes, and Relative Risks
175 (natural logarithm) for safety outcomes.

176 2.6. Data analysis and reporting

177 The aims of this study were to: (1) update our previous study by performing
178 conventional (frequentist) random-effect meta-analysis with DerSimonian and Laird
179 method and (2) conduct frequentist and Bayesian NMA to assess comparative efficacy
180 and safety of anti-amyloid antibodies. The results obtained from different methods were
181 then compared for each outcome. In order to properly manage the differences among
182 multiple subgroups within the primary studies with respect to differences in drug doses,
183 biomarker findings and other patient characteristics, we followed the recommendation
184 to treat each subgroup as a separate study [21].

185 2.6.1. Frequentist NMA

186 Frequentist random-effects NMA was conducted using the *netmeta* package [22] in R
187 (version 4.2.2). The values of I^2 statistic and their statistical significance were evaluated
188 for overall heterogeneity or inconsistency in the network, which was then split into
189 within-design (i.e., heterogeneity) and between-design (i.e., inconsistency) and the
190 statistical significance of Q-values was assessed. For each outcome, the results are
191 presented in the form of league tables and treatment rankings calculated as P-scores by
192 using *netrank* function from *netmeta* package [22-24]. P-score of 1 is the best possible
193 score and 0 is the worst. Further information on this ranking can be found in the Method
194 section of the Supplement.

195 2.6.2. Bayesian NMA

196 Bayesian random-effect NMA was performed by using *rjags* and *gemtc* packages in R
197 [25] and JAGS 4.3.0 software [26] that allow automating NMA in the Bayesian
198 hierarchical model framework with uninformative priors and Markov chain Monte
199 Carlo (MCMC) estimation. The number of tuning iterations was set at 10 and the

200 number of simulation iterations at 100,000. Convergence of the MCMC model was
201 assessed through Gelman-Rubin plots and the potential scale reduction factor. Trace and
202 density plots were then applied to determine the best parameters of the number of burn-
203 in iterations, the number actual simulation iterations, and thinning parameter. For each
204 outcome, the results are presented in the form of league tables and rankings are
205 calculated in the form of SUCRA score (surface under the cumulative ranking curve).
206 This score can be interpreted as the estimated proportion of treatments worse than the
207 treatment of interest ($0\% \leq \text{SUCRA} \leq 100\%$). Values of SUCRA close to 1 (100%)
208 indicate the higher likelihood that specific treatment is the best. To calculate the
209 SUCRA scores in R, we used the *sucra* function from *dmetar* package in R [23]. Further
210 information on SUCRA score can be found in the Supplement.

211 2.6.3. Influence and Sensitivity Analyses

212 Influence analyses were done to detect outliers and influential studies in the network.
213 These included heterogeneity assessments using Baujat plots, Influence diagnostics, and
214 Leave-One-Out meta-analysis. Then, graphic display of study heterogeneity (GOSH)
215 plot analysis was used to examine the results of an equal-effects model in the
216 combinatorial meta-analyses with all possible subsets of studies [23, 27]. GOSH plot
217 diagnostics employs three unsupervised machine learning algorithms: (1) k-means
218 clustering based on vector quantization [28], (2) density-based spatial clustering of
219 applications with noise (DBSCAN) [29], and (3) probabilistic Gaussian mixture model
220 (GMM) [30]. Sensitivity analyses were then performed by rerunning frequentist and
221 Bayesian NMAs upon excluding influential observations.

222 2.6.4. Rank Robustness

223 To quantify the influence a study had on the SUCRA based treatment ranks, the

224 robustness of ranks was evaluated within the sensitivity analyses, upon excluding all
225 influential studies and during Leave-One-Out meta-analysis. The influence of study
226 elimination on the ranks in the Bayesian NMA was estimated by using quadratic
227 weighted Cohen's kappa [31], as recently proposed by Daly and colleagues [32]. This
228 allowed us to measure the agreement between the ranks produced with the complete
229 dataset and the ranks produced from sub-datasets generated upon removing influential
230 studies and each study at a time. The magnitude of Cohen's kappa is usually categorised
231 into levels of agreement for interpretation (eg, poor (<0%), slight (0%–20%), fair
232 (21%–40%), moderate (41%–60%), substantial (61%–80%) and almost perfect (81%–
233 100%) agreement). However, it should be added that a kappa is not an index of accuracy
234 but of agreement beyond chance [33], and that the uncertainty of rankings and other
235 study characteristics that are context-dependent should be incorporated into the
236 interpretation of the results, as a kappa value of 90%, for example, is not always
237 indicative of almost perfect rank agreement [32]. The calculations were performed by
238 using `cohen.kappa` function available in *psych* package in R [34]. In addition, the
239 robustness of each rank was examined by determining how many (%) studies did not
240 change any treatment rank (%), as well as how many (%) studies displaced treatment
241 ranks.

242 [2.6.5. Rank Uncertainty](#)

243 The uncertainty of the SUCRA ranks in Bayesian NMA is visualized by rankograms and
244 tables that show the distribution of the ranking probabilities [35-37], as suggested by the
245 PRISMA–NMA guideline and GRADE working group. Furthermore, we measured
246 Shannon's information entropy, as recommended recently by Wu et al (2021) [38]. The
247 authors suggested very intuitive way to estimate uncertainty associated with ranking
248 probabilities in NMA by applying Shannon's information entropy formula to obtain a

249 normalised entropy score. Their work showed that the normalised entropy score gives a
250 more accurate assessment of ranking uncertainty and does not depend on the number of
251 analysed treatments. This approach can be used to compare the uncertainty of treatment
252 rankings within a NMA and also between different NMAs, which is crucial for
253 interpretation of results [38].

254 For a NMA, the most precise scenario would be that we are absolutely certain in the
255 ranking of treatments in our network. Therefore, each treatment would have 100%
256 probability of being in one ranking position and 0% probability for the other positions
257 (peaked distribution). Under this scenario, the entropy is zero bit, and Shannon's
258 normalized entropy equals zero. On the other hand, the normalized entropy reaches 1 in
259 the least precise scenario when the ranking probabilities are the same for each rank (flat
260 distribution). In this work, we regarded as "low" those values of normalized entropy
261 that were equal or lower than 0.25 and "moderate" those between 0.25 and 0.5. Larger
262 values were considered as conveying "high" or "substantial" uncertainty or normalized
263 entropy associated with ranking probabilities. Further information on normalized
264 entropy score, and the code for calculation are located in the Supplement.

265 [2.6.6. Clinical relevance](#)

266 Finally, the clinical relevance of the statistically significant results (compared to
267 Placebo) from the frequentist main NMA and sensitivity analyses was assessed by
268 calculating the "Number Needed to Treat" (NNT) or "Number Needed to Harm" (NNH)
269 from the treatment's pooled effect size (Hedges g') by Area Under the Curve (AUC)
270 method [39]. Here, AUC is defined as the probability that a patient receiving the
271 treatment has an outcome preferable to one receiving Placebo. Lower NNT and higher
272 NNH values are associated with a more favorable benefit/risk profile. The best NNT is
273 1, indicating that for every patient treated one got better, while $NNT > 1$, means that the

274 fewer people will benefit the treatment [23]. Similarly, NNH of 1 means that for every
275 patient treated one got harmed, while $NNH > 1$ suggests that fewer people will be
276 harmed.

277 **2. RESULTS**

278 The results obtained from the conventional (frequentist) meta-analysis with
279 DerSimonian-Laird estimator were very similar to the Frequentist and Bayesian NMA
280 results. Due to word limitation and large consistency in the findings, this manuscript
281 will only refer to the results obtained from the NMAs. Conventional meta-analysis will
282 be reproducible online in AlzMeta.app (<https://alzmetaapp.shinyapps.io/alzmeta/>), upon
283 publication of this work.

284 The meta-analysis included 16 studies (11 ClinicalTrials.gov registries), considering
285 different subgroups within certain RCTs as separate studies, according to our
286 methodology [21]. Bapineuzumab was tested in six studies [40, 41], Solanezumab in
287 three studies [42, 43] Aducanumab in four studies ([44]), Donanemab in two studies
288 (low/medium-tau and high-tau populations from TRAILBLAZER-ALZ 2 [8], and
289 Lecanemab in one study [45]. The details concerning the quality of reports from
290 primary studies (modified Jadad scale) can be found in the Supplement (Supplementary
291 Tables A1 and A2), as well as the number of records identified, included and excluded,
292 and the overall reasons for exclusions (PRISMA Flow Diagram, Supplement), along
293 with PRISMA Checklist (Supplement). Specific reasons for excluding each primary
294 study are listed in Supplementary Tables A3 and A4 (Supplement).

295 In total, 16,971 patients were analyzed in the primary studies, 8,447 (mean age 71,43
296 years) received Placebo, while 8,524 (mean age 71,87) were passively immunized
297 against A β . All treatments lasted either 76 or 78 weeks across multiple centers, with the

298 average number of trial sites 212.27 (median 198). Acetylcholinesterase inhibitors and
299 memantine were allowed during the trials, alone or in combination. Additional baseline
300 participant characteristics and the inclusion criteria for primary studies are documented
301 in the Supplement (Table A5 and A6).

302 Statistical power was 100% assuming moderate between-studies heterogeneity and any
303 possible effect size (small, medium, large) (Fig. S1, Supplement). Regarding risk of
304 bias, all studies were considered as either “low risk” or raising “some concerns” (Fig.
305 S2a,b and Table A7; Supplement). For each outcome, certainty of evidence is rated low
306 or very low within each comparison, according to GRADE approach (Table A8;
307 Supplement), with low number of studies being the most important reason for
308 downgrading the certainty.

309 [2.1. Primary outcomes](#)

310 Frequentist NMA results showed that Donanemab, Lecanemab and Aducanumab were
311 significantly more effective than Placebo on ADAS-Cog (cognitive scale, Table 1A).
312 Bapineuzumab and Solanezumab had no favorable effects, with frequentist (but not
313 Bayesian) NMA results suggesting that Placebo was superior to Solanezumab on this
314 scale. In frequentist framework, Donanemab was significantly more effective than
315 Bapineuzumab and Solanezumab and had a double effect size compared to that
316 achieved with Aducanumab and Lecanemab. Bayesian NMA results revealed strong
317 evidence that Aducanumab and Donanemab improved ADAS-Cog and the lack of
318 convincing support for the effects of Lecanemab. Furthermore, in Bayesian framework,
319 we found strong indication that Donanemab outperformed Bapineuzumab, but not
320 Solanezumab, which was clear from the credible intervals including zero (Table 1B).

321 Even better results were obtained on cognitive and functional (global) scale CDR-SB
322 (Table 1C), where Donanemab, Lecanemab and Aducanumab were significantly more
323 effective than Placebo. Donanemab was more effective than other interventions with
324 nearly three times greater effect size than that of the approved antibodies. Aducanumab
325 and Lecanemab were significantly more effective than Bapineuzumab, however, with
326 the confidence intervals very close to zero. Bayesian NMA produced very similar
327 results as frequentist NMA on CDR-SB, with the exception of three indirect
328 comparisons (Donanemab vs. Lecanemab; Lecanemab vs Bapineuzumab and
329 Aducanumab vs Bapineuzumab), where credible intervals included zero, suggesting the
330 lack of strong evidence for the non-zero effect in these comparisons (Table 1D).

331 Frequentist NMA results for MMSE (cognitive) score revealed weaker effects of all
332 drugs in the network, with Donanemab and Solanezumab significantly more effective
333 than Placebo (MMSE results for Lecanemab were not reported). These results were
334 fully validated by Bayesian NMA (Table S1; Supplement).

335 Heterogeneity was not significant for any primary outcome, and inconsistency was not
336 assessed because the network did not have a closed loop. We found no evidence of
337 publication bias, with the Egger's test p-values 0.661, 0.8462, and 0.8929 for ADAS-
338 Cog, CDR-SB, and MMSE, respectively. Funnel plots were symmetric for all three
339 outcomes (Fig. S3a-c in the Supplement), and the imputation of potentially "missing
340 studies" did not change results.

341 Across all primary outcomes, Donanemab was the best-ranked antibody according to
342 both frequentist P-scores [24] and SUCRA scores [35] in Bayesian framework. On
343 CDR-SB, Donanemab obtained the maximum scores (near 1) and very low Shannon's
344 normalized (information) entropy score [38]. This suggests high certainty in the ranking
345 probabilities for Donanemab on CDR-SB. In contrast, on ADAS-Cog and MMSE,

346 SUCRA scores for Donanemab were 0.93 and 0.89 respectively, with moderate
347 uncertainty in the ranking probabilities (Fig. 1C and Table P1 in the Supplement).
348 Donanemab produced clinically the most relevant effects on all primary outcomes in
349 combined low/medium population and high-*tau* population, with NNT = 9 on ADAS-
350 Cog, 6 on CDR-SB and 13 on MMSE. When Donanemab was compared to
351 Aducanumab and Lecanemab, the most striking difference was observed on CDR-SB,
352 with NNT values almost two times higher for Lecanemab (13) and almost three times
353 higher for Aducanumab (18). On ADAS-Cog, Lecanemab and Aducanumab also had
354 high NNT (15). Finally, Solanezumab had high NNT (20) on MMSE, suggesting low
355 clinical relevance of cognitive effects of this drug.

356 2.2. Secondary outcomes

357 2.2.1. Biomarkers of A β and *tau*

358 Both Frequentist and Bayesian NMA results showed that Lecanemab, Donanemab, and
359 Aducanumab significantly improved clearance of A β deposits from brain when
360 compared to Placebo (Table S2.1; Supplement). Lecanemab demonstrated the greatest
361 amyloid-clearance activity on PET and obtained P- and SUCRA scores very close to 1,
362 with high certainty (low information entropy) associated with the ranking probabilities
363 (Table P1 in the Supplement). However, despite large effect-sizes, these data should be
364 interpreted with caution, as substantial heterogeneity was found between the studies and
365 the number of analyzed trials/patients was small. While frequentist NMA results
366 indicated significant differences between Lecanemab and other antibodies, Bayesian
367 NMA results revealed a lack of strong evidence that Lecanemab outperformed
368 Aducanumab and Donanemab on this outcome.

369 Frequentist NMA results revealed that Aducanumab and Lecanemab were significantly
370 superior to Placebo on CSF biomarkers of $A\beta_{1-42}$ and p-181-*tau* (data for Donanemab
371 were not reported). Aducanumab significantly outperformed Solanezumab on CSF $A\beta$
372 (Table S2.1) and Bapineuzumab on CSF p-*tau* (Table S2.2.). Bayesian NMA validated
373 the relative effect of Aducanumab on both CSF $A\beta$ and p-*tau*, providing strong evidence
374 in favor of positive effects of Aducanumab on these outcomes when compared to
375 Placebo, but not when compared to other drugs. Furthermore, Bayesian NMA results
376 revealed an absence of strong evidence that Lecanemab was superior to Placebo on CSF
377 $A\beta$ and p-*tau*. Aducanumab was ranked the best treatment on both CSF biomarkers of
378 $A\beta$ and p-*tau*, with SUCRA scores 0.86 and 0.89, respectively, and moderate
379 uncertainty in the ranking probabilities (Table P1; Supplement).

380 As in case with amyloid PET data, these findings should be interpreted carefully, since
381 CSF biomarker analyses were available only for limited number of patients in few
382 studies, and considerable heterogeneity was found for CSF $A\beta$ ($I^2 = 68.9\%$ [20.0%;
383 87.9%]; p-value of Q statistics = 0.012), indicating that the true effect size may vary
384 across studies. Heterogeneity was not significant for CSF p-*tau* (Table S2.2;
385 Supplement), while the analysis of plasma *tau* biomarkers had to be dropped because of
386 extreme heterogeneity ($I^2 = 99\%$ [98.5%; 99.3%], $p < 0.0001$) and differences in
387 measured biomarker outcome.

388 All biomarker effects were clinically relevant for Aducanumab, Lecanemab and
389 Donanemab, with NNT values of 1 for the effects on amyloid burden (PET data),
390 indicating that the favorable effect on amyloid removal was accomplished for each
391 treated patient, with each of these antibodies. NNTs for CSF outcomes were also very
392 small (clinically relevant) for Aducanumab and Lecanemab, with NNT = 2 on CSF $A\beta_{1-42}$
393 for both antibodies. As for CSF-p-*tau*, the NNT of Aducanumab (3) was somewhat

394 lower than that of Lecanemab (4), reflecting slightly greater effect of Aducanumab on p-
395 *tau*.

396 2.2.2. Safety outcomes

397 Anti-amyloid antibodies did not increase the risk of serious adverse events (Table S3.1;
398 Supplement). Nevertheless, both the frequentist and Bayesian NMA results provide
399 substantial evidence that high-clearance antibodies (Donanemab, Lecanemab and
400 Aducanumab) were less tolerable than Placebo and Solanezumab since these drugs
401 greatly increased the risk of treatment discontinuation due to adverse events (Table 2).
402 Donanemab was the least tolerable antibody in the network, significantly worse than
403 Bapineuzumab and Aducanumab. Frequentist NMA results revealed considerable risks
404 of discontinuation for Bapineuzumab when compared to Placebo, however, this was not
405 corroborated in Bayesian analysis.

406 Frequentist and Bayesian NMA results provide compelling evidence that Solanezumab
407 was the most tolerable and the safest antibody in terms of ARIAs (Fig. 1C, Table 2).
408 Other antibodies substantially increased the total risk of ARIA-E and the risk of cerebral
409 microhemorrhages (ARIA-H), when compared to Placebo and Solanezumab.
410 Furthermore, Aducanumab, Donanemab and Lecanemab significantly increased the total
411 risk of ARIA-H (including all the events of cerebral superficial siderosis, micro- and
412 macrohemorrhages), when compared to Solanezumab and Placebo (Tables S3.2 and
413 S3.3 in the Supplement)

414 Frequentist NMA on additional safety outcomes showed that Aducanumab significantly
415 increased the risk of headaches (Table S3.4) and falls (Table S3.5; Supplement) when
416 compared to Placebo, however, with confidence intervals very close to zero and
417 substantial heterogeneity between the studies (see 3.6). Further Bayesian NMA revealed

418 lack of conclusive support for these comparisons. Similarly, in Bayesian framework,
419 there was an absence of strong evidence that Donanemab increased the risk of nausea in
420 high-*tau* population, when compared to Placebo, whereas frequentist NMA would
421 suggest significant difference in risk (Table S3.6; Supplement).

422 Further analyses showed that anti-A β did not raise the risks of fatigue, dizziness,
423 arthralgia, back pain, cardiac disorders, diarrhea, nor urinary infections more than
424 Placebo (Table S3.7-S3.16; Supplement). In addition, we found no improved risk of
425 syncope, nasopharyngitis and upper respiratory infections, however, these events were
426 not reported for Donanemab and Lecanemab, antibodies tested in clinical trials affected
427 by COVID-19 pandemic (see 2.4 section). Heterogeneity was not significant for
428 secondary safety outcomes with the exception of risks of headaches and cardiac
429 disorders.

430 The risks of treatment discontinuation due to adverse events, total ARIA-E, cerebral
431 microhemorrhages, and total ARIA-H were all clinically relevant, with NNH values of
432 1, which correspond to the situation in which for each patient treated, one exhibited
433 treatment-emergent adverse reaction.

434 [2.3. Tertiary outcomes](#)

435 Frequentist NMA of ADCS-MCI-ADL showed that Aducanumab (NNT = 12) and
436 Lecanemab (NNT = 10) significantly slowed functional decline by small effect size,
437 when compared to Placebo. Bayesian NMA results validated these results for
438 Aducanumab, but not for Lecanemab (Table S4.1). The heterogeneity was not
439 significant; however, these results should be interpreted with care since they are based
440 on four studies with Aducanumab and one study testing Lecanemab. Furthermore,

441 credible intervals for the relative effects of Aducanumab (vs. Placebo) were very close
442 to zero, indicating uncertainty in the evidence (also, see 2.4).

443 Both frequentist and Bayesian NMA results gave strong support for increased risk of
444 superficial siderosis (ARIA-H) in patients treated with Donanemab and Aducanumab,
445 when compared to those in control group, while the results for Lecanemab were not
446 conclusive. Whereas frequentist NMA results suggested significantly greater risk ($p =$
447 0.001) of superficial siderosis in patients treated with Lecanemab (compared to
448 Placebo), Bayesian NMA gave an estimate with overlapping credible intervals,
449 suggesting the lack of strong evidence for this comparison (Table S4.2). Even so, the
450 risk of superficial siderosis was clinically relevant for Donanemab, Aducanumab and
451 Lecanemab, with NNH values of 1 (the least favorable).

452 The risk of infusion-related reactions was clinically meaningful (NNH = 1) for
453 Donanemab and Lecanemab (not reported for other drugs), significantly higher ($p <$
454 0.0001) when compared to Placebo. Both the Frequentist and Bayesian NMA results
455 provided strong evidence that Donanemab increased the risks of infusion-related
456 reactions, with almost two times greater risk than that observed in Lecanemab study
457 (Table S4.3 in the Supplement).

458 Having in mind the limited information regarding ARIA-E and ARIA-H depending on
459 the APOE status, frequentist and Bayesian NMAs strongly indicate that all antibodies,
460 except Solanezumab, significantly increased the risk of ARIA-E in both APOE- $\epsilon 4$
461 carriers and non-carriers. The results were not that convincing for ARIA-H, where
462 smaller risks were observed and frequentist NMA results suggested significant increase
463 in risk for Donanemab and Lecanemab in both carriers and non-carriers, which was not
464 validated in Bayesian framework. Due to inconclusive findings on ARIAs depending on
465 APOE status, we were unable to determine whether the risk of ARIAs increases

466 substantially with $\epsilon 4$ allele dose (homozygotes vs. heterozygotes). Overall, the risk of
467 ARIAs was clinically relevant (NNH = 1) independently of APOE status, and the risk
468 generally increased with each $\epsilon 4$ allele (Tables S4.4-S4.7 in the Supplement).
469 Heterogeneity between the studies could not be assessed due to small numbers of
470 studies.

471 2.4. Sensitivity analyses

472 Donanemab consistently showed greater cognitive effects in patients with low/medium
473 *tau* load than in high-*tau* population (assessed by PET imaging), and removing this
474 study had a big impact on the results, similarly to the greater effect observed for high-
475 dose Aducanumab in controversial EMERGE study, as found in our previous meta-
476 analysis [46] (Tables S5 and S7 in the Supplement). For this reason, removing these two
477 studies from the NMAs yielded much lower relative effects of these drugs on cognitive
478 measures (when compared to Placebo), with p-values of 0.0502 and 0.5334 on ADAS-
479 Cog and 0.1419 and 0.0249 on CDR-SB for Aducanumab and Donanemab, respectively.
480 Likewise, excluding study in low/medium-*tau* population greatly reduced relative
481 effects of Donanemab on MMSE ($p = 0.3430$), further emphasizing favorable effects of
482 this drug in patients with low/medium *tau*, but not high *tau* burden. Bayesian NMA
483 validated these findings, strongly indicating that CDR-SB score was improved by
484 Donanemab (even in high *tau* population), but not by Aducanumab, after excluding
485 controversial EMERGE study. There was lack of strong evidence for improving both
486 ADAS-Cog and MMSE scores neither by Aducanumab nor by Donanemab after
487 removing their influential studies. Furthermore, frequentist NMA results of amyloid
488 PET data revealed that both Donanemab and Aducanumab significantly removed
489 amyloid brain burden (Table S7.4), when compared to Placebo. Conversely, the results
490 from Bayesian NMA provided strong evidence supporting the efficacy of Aducanumab,

491 but not Donanemab. In other words, comparatively higher effect of Donanemab was
492 obtained in low/medium-*tau* population, and removing this measurement increased the
493 uncertainty of the evidence, shifting the credible intervals for the relative efficacy of
494 Donanemab from (-6.4251; -2.0495) towards (-9.3948; 1.7020). Finally, frequentist
495 NMA results of ADCS-ADL-MCI showed that removing high-dose Aducanumab
496 EMERGE trial did not impact the relative effect of this drug when compared to Placebo
497 (Table S7.7). Bayesian NMA weakly favored Aducanumab over Placebo, however, with
498 credible intervals (-0.014; 0.245) including zero. In summary, significant dose-
499 dependent efficacy of Aducanumab was evident only in EMERGE study, and the
500 relative effects of Donanemab in the main study were largely due to its beneficial effects
501 on AD patients with low/medium *tau* load.

502 Removing the influential studies of Donanemab and Aducanumab reduced ranking
503 scores of these drugs in the network (Figure 1F). On MMSE score, for example,
504 Solanezumab was the best-ranked drug with SUCRA score 0.81, followed by
505 Donanemab in high-*tau* population (0.71). Lecanemab obtained the best score (0.79) on
506 ADAS-Cog, followed by Aducanumab (0.69), and Solanezumab (0.58). Donanemab in
507 high-*tau* population was the best-ranked treatment only on CDR-SB score (0.91) and
508 fourth-ranked on ADAS-Cog (0.3). Additionally, Lecanemab was better ranked
509 treatment than Aducanumab on CSF-A β_{1-42} and ADCS-ADL-MCI, upon excluding the
510 influential effects from high-dose Aducanumab in EMERGE study. For efficacy
511 outcomes, the most robust SUCRA scores were obtained for CDR-SB, with 92% of
512 agreement with the main analysis (median of observed weighted Cohen's kappa values)
513 and 8 (50%) studies that did not change any treatment rank. The robustness of SUCRA
514 scores further improved upon excluding influential Aducanumab and Donanemab
515 studies, however, the uncertainty in ranking probabilities also increased. The least

516 robust SUCRA scores were found for amyloid burden on PET, which was a direct
517 consequence of small number of studies. Across all efficacy outcomes, the treatments
518 moved only one or very rarely two ranks up or down upon excluding each study at a
519 time. Further details on ranking probabilities and their sensitivity analysis can be found
520 in the Supplement (Tables P1-P4; Supplement).

521 Clinical relevance of the cognitive and functional effects of Donanemab and
522 Aducanumab decreased remarkably upon removing their influential studies. while the
523 biomarker effects on $A\beta$ and *p-tau* remained clinically relevant with small NNTs. For
524 Donanemab, the NNT was 28 for ADAS-Cog and 18 for MMSE, while NNT for CDR-
525 SB did not change drastically (NNT = 8), when compared to the main analysis (NNT =
526 6). In other words, the beneficial effects of Donanemab on cognitive and functional
527 CDR-SB score in patients with high-*tau* load were still clinically relevant, even if the
528 effects were greater in those with low/medium *tau*. Aducanumab had higher (less
529 favorable) NNTs in sensitivity analysis for ADAS-Cog (21) and CDR-SB (19).

530 The results of sensitivity analyses on safety outcomes were largely consistent with the
531 main analysis. SUCRA scores for treatment discontinuation, risks of ARIA-E and
532 ARIA-H were robust to study exclusions, with median weighted Cohen's kappa values
533 above 89%, indicating almost perfect agreement in study ranks between main and
534 sensitivity analysis (Tables P1-P4 in the Supplement). Additionally, there was small
535 statistically significant increase in the risk of headaches for all high-clearance
536 antibodies: Aducanumab ($p = 0.0001$), Donanemab ($p = 0.0085$), and Lecanemab ($p =$
537 0.0323), when compared to Placebo (Table S2.9), after excluding two low-dose
538 Bapineuzumab studies [41]. These two studies were the source of high heterogeneity in
539 the main analysis, which was detected by Q statistics ($p < 0.0001$) and Gaussian
540 Mixture Model in GOSH plot analysis (Table S8.1 and Figure S11 in the Supplement).

541 Bayesian NMA validated frequentist NMA results for Aducanumab and Donanemab,
542 suggesting that these antibodies slightly increased the risk of headaches when compared
543 to Placebo and Solanezumab. On the other hand, there was no conclusive evidence that
544 Lecanemab elevated risk of headaches, when compared to Placebo. Additionally, the
545 statistical significance for the risk of falls in Aducanumab studies dropped ($p = 0.8785$)
546 once we removed ENGAGE studies (Table S8.2 in the Supplement), suggesting that
547 these studies contributed to the observed high risk of falls in the main analysis.

548 [2.5. The Impact of COVID-19 pandemic](#)

549 Lecanemab and Donanemab RCTs included in this study were affected by global
550 outbreak of coronavirus – an infectious disease caused by the severe acute respiratory
551 syndrome. COVID-19 was the most commonly reported adverse event across treatment
552 groups in Donanemab studies, occurring in ~17% of patients in all arms. The impact of
553 pandemic was less severe on Lecanemab study, with the incidence of COVID-19 around
554 7% in experimental and control groups. The pandemic also caused delays in study visits
555 and assessments in both trials and other difficulties summarized in the Supplement
556 (Table S6).

557

558 3. DISCUSSION

559 This study reveals a complex pattern of evidence supporting the efficacy of
560 Aducanumab, Lecanemab and Donanemab across multiple cognitive, functional and
561 biomarker outcomes. At the same time, the risks of cerebral edema and microbleeding
562 were substantial, suggesting greater harm than benefit. The similarities and the
563 differences between frequentist and Bayesian NMA results presented here strongly
564 emphasize the need for further studies with larger sample sizes and the need to study
565 these antibodies in earlier stages of AD and MCI, and in younger patients with lower A β
566 and *tau* load.

567 The inconsistencies between ENGAGE and EMERGE trials revealed here and in our
568 previous study [46] have been discussed in literature ever since these studies were
569 published, as they caused a lot of controversy within the FDA and in the broader
570 community [47]. The Aducanumab provider, in partial collaboration with the FDA,
571 provided two possible explanations for these discrepancies: (1) fewer numbers of
572 patients in the ENGAGE trial receiving higher doses of drug, and (2) ENGAGE trials
573 had more outliers, and removing these outliers in sensitivity analysis gave more
574 compatible dose-dependent beneficial effects [48, 49]. Despite these inconsistencies,
575 Aducanumab received accelerated approval in the U.S. Afterwards, the FDA granted
576 breakthrough status to Lecanemab and Donanemab, in order to speed up their approval
577 [50], which finally resulted in full FDA approval of Lecanemab in the U.S. and more
578 recently in Japan. Furthermore, Lecanemab is currently undergoing regulatory review in
579 the European Union (EU), United Kingdom, South Korea, and Canada, while
580 Donanemab is reviewed in the EU and the U.S [7]. Recent exploratory *post hoc*
581 modeling based on phase III RCT of Donanemab suggested that A β levels in immunized
582 patients would remain below the positivity threshold for almost 4 years without the

583 treatment. The authors also revealed that greater plaque removal by Donanemab was
584 associated with slower progression of *tau* PET and slower clinical decline, but only in
585 APOE- ϵ 4 carriers [51]. In our study, all high-clearance antibodies (Aducanumab,
586 Donanemab and Lecanemab) produced substantial A β and *tau* biomarker effects.
587 Donanemab was the best-ranked treatment across all cognitive measures, while
588 Lecanemab was the most efficient at reducing amyloid burden on PET. Donanemab
589 consistently showed better cognitive, functional and biomarker effects in patients with
590 low/medium *tau* load, compared to those with high *tau*. Donanemab produced clinically
591 relevant cognitive effects in low/medium-*tau* population on all three primary outcomes,
592 suggesting greater benefit of Donanemab in earlier AD stages.

593 Low tolerance and high risk of ARIAs looms over high-clearance anti-A β antibodies.
594 Our results showed that Donanemab was significantly less safe and less tolerable than
595 Placebo and Solanezumab, with similar risks of treatment discontinuation and ARIAs as
596 in Aducanumab and Lecanemab studies. Both frequentist and Bayesian NMA results
597 from this study imply that the risk of ARIAs associated with the treatment may
598 outweigh the benefits, independently of APOE status. In summary, our results
599 demonstrate modest cognitive and substantial biomarker effects of Donanemab,
600 Lecanemab and Aducanumab, yet with significant trade-off in the form of dose-
601 dependent cerebral vasogenic edema (ARIA-E), bleeding and hemosiderin deposits
602 (ARIA-H), which may or may not be symptomatic or serious adverse event. These risks
603 require careful dose titration and close patient monitoring on Magnetic Resonance
604 Imaging (MRI), which increases the costs of already expensive treatments [52]. In our
605 view, additional and comparative studies of these drugs with vascular biomarkers are
606 needed to evaluate their risk/benefit profile, as well as more anticipatory policies from
607 regulatory bodies if we were to avoid new controversies as those seen with the FDA

608 approval of Aducanumab. Moreover, personalized approach and safer multi-target
609 therapies in combination with novel biomarkers is expected to effectively mitigate the
610 AD progression [53].

611 Different mechanisms of actions of these drugs outlined in the Introduction certainly
612 contribute to the differences in their effects and heterogeneity in our networks.

613 However, there is no doubt that Aducanumab, Lecanemab and Donanemab achieved
614 great biomarker results and modest cognitive effects in 18 months. Therefore,
615 comparative and longitudinal studies are required to fully disclose the impact of anti-A β
616 immunotherapy over longer periods of time. Future trials would need to compare the
617 effects of anti-A β antibodies in multiple stages of AD and in different populations, as
618 inter-individual differences and comorbidities may be crucial to prevent ARIAs and
619 select patients that could benefit from anti-A β immunotherapy. The inflammation and
620 vascular comorbidities are of special concern, as blood-brain barrier disruption and A β
621 deposition are central to the development of both AD (with A β deposited in brain
622 parenchyma) and cerebral amyloid angiopathy (with A β deposition on blood vessel
623 walls). These two pathologies often co-occur and APOE- ϵ 4 status is known to worsen
624 both of them. Risk factors for ARIA are dose-dependent and include receiving anti-A β
625 immunotherapy, anticoagulant and thrombolytic medication, carrying APOE-e4 allele,
626 and history of brain microhemorrhages and stroke [54, 55].

627 The important limitations of this study are (1) COVID-19 impacted Donanemab and
628 Lecanemab trials, (2) the patients in Donanemab study were stratified by their level of
629 the brain protein *tau*; (3) the effects of antibodies on plasma *tau* biomarkers could not be
630 evaluated due to excessive between-study heterogeneity and different measures used in
631 primary studies; (4) potential bias might be introduced due to unmasking, since ARIAs
632 were the most frequent treatment-emergent adverse events of all drugs except

633 Solanezumab, and (5) direct comparison between the antibodies has not been done yet,
634 and our NMA only included studies where treatments were compared with Placebo.
635 NMA is a statistical method to compare more than two interventions simultaneously by
636 combining direct and indirect evidence. When head-to-head trials are not available, as
637 often happens in many research fields, indirect comparisons can be made via a common
638 control arm (Placebo in our case) [56]. Therefore, this study should be viewed as a way
639 to summarize the results from previous phase III RCTs and provide comparative
640 assessment of these drugs based on existing data, until further evidence comes along.
641 The same should be emphasized for the efficacy and safety rankings and their
642 uncertainty measures, as clinical decision making should never be based solely on one
643 or two measures or studies.

644

645 **4. RESEARCH IN CONTEXT**

- 646 1. Systematic review and meta-analysis: Frequentist and Bayesian network meta-
647 analyses (NMAs) were performed to compare the efficacy and safety of anti-
648 amyloid antibodies in phase III randomized placebo-controlled trials in sporadic
649 AD.
- 650 2. Interpretation: Aducanumab, Lecanemab and Donanemab achieved substantial
651 biomarker results and small cognitive effects, and were less tolerable than
652 Placebo and Solanezumab, significantly increasing the risk of ARIAs.
653 Donanemab was the most efficient antibody on the cognitive measures, but it was
654 also the least tolerable, significantly less tolerable than Placebo, Solanezumab,
655 Bapineuzumab and Aducanumab.
- 656 3. Future directions: Studying these drugs in earlier stages and in younger patients
657 with lower A β and *tau* burden might yield better outcome than that observed in

658 phase III trials. Since ARIAs are the major adverse events limiting the broad use
659 of anti-amyloid antibodies, prevention of vascular damage, better biomarkers and
660 safer, personalized treatment options are needed to improve AD treatment.

661

662

663 **ACKNOWLEDGEMENTS**

664 This work was supported by grants PID2020–115823-GB100 funded by
665 MCIN/AEI/10.13039/501100011033, and SBPLY/21/180501/000150 funded by
666 JCCM/ERDF - A way of making Europe, and 2022-GRIN-34354 grant by
667 UCLM/ERDF intramural funding to LJ-D and JDN-L.

668 **AUTHORS' CONTRIBUTIONS**

669 JDNL and LJD were responsible for the initial conceptualization, funding acquisition,
670 supervision, and project administration; DJ was responsible for Data curation, Software,
671 Formal analysis, visualization and writing the original draft; SD, JDNL and LJD did the
672 writing – review and editing. All authors read and approved the final manuscript.

673 **CONFLICT OF INTEREST STATEMENT**

674 All authors declare that they have no conflicts of interest. The protocol for this study
675 was not previously registered. Author disclosures are available in the Supporting
676 Information.

677

678 **REFERENCES**

679

- 680 [1] Holtzman DM, Morris JC, Goate AM. Alzheimer's disease: the challenge of the second
681 century. *Sci Transl Med*. 2011;3:77sr1.
- 682 [2] Wessels AM, Dennehy EB, Dowsett SA, Dickson SP, Hendrix SB. Meaningful Clinical
683 Changes in Alzheimer Disease Measured With the iADRS and Illustrated Using the
684 Donanemab TRAILBLAZER-ALZ Study Findings. *Neurol Clin Pract*. 2023;13:e200127.
- 685 [3] Söderberg L, Johannesson M, Nygren P, Laudon H, Eriksson F, Osswald G, et al.
686 Lecanemab, Aducanumab, and Gantenerumab - Binding Profiles to Different Forms of
687 Amyloid-Beta Might Explain Efficacy and Side Effects in Clinical Trials for Alzheimer's
688 Disease. *Neurotherapeutics*. 2023;20:195-206.
- 689 [4] Plotkin SS, Cashman NR. Passive immunotherapies targeting A β and tau in
690 Alzheimer's disease. *Neurobiol Dis*. 2020;144:105010.
- 691 [5] Prins ND, Scheltens P. Treating Alzheimer's disease with monoclonal antibodies:
692 current status and outlook for the future. *Alzheimers Res Ther*. 2013;5:56.

- 693 [6] Rabinovici GD, La Joie R. Amyloid-Targeting Monoclonal Antibodies for Alzheimer
694 Disease. *Jama*. 2023;330:507-9.
- 695 [7] Cummings J, Osse AML, Cammann D, Powell J, Chen J. Anti-Amyloid Monoclonal
696 Antibodies for the Treatment of Alzheimer's Disease. *BioDrugs*. 2024;38:5-22.
- 697 [8] Sims JR, Zimmer JA, Evans CD, Lu M, Ardayfio P, Sparks J, et al. Donanemab in Early
698 Symptomatic Alzheimer Disease: The TRAILBLAZER-ALZ 2 Randomized Clinical Trial.
699 *Jama*. 2023;330:512-27.
- 700 [9] Digma LA, Winer JR, Greicius MD. Substantial Doubt Remains about the Efficacy of
701 Anti-Amyloid Antibodies. *J Alzheimers Dis*. 2024;97:567-72.
- 702 [10] Lacorte E, Ancidoni A, Zaccaria V, Remoli G, Tariciotti L, Bellomo G, et al. Safety and
703 Efficacy of Monoclonal Antibodies for Alzheimer's Disease: A Systematic Review and
704 Meta-Analysis of Published and Unpublished Clinical Trials. *J Alzheimers Dis*.
705 2022;87:101-29.
- 706 [11] Jeremic D, Navarro-López JD, Jiménez-Díaz L. Efficacy and safety of anti-amyloid- β
707 monoclonal antibodies in current Alzheimer's disease phase III clinical trials: A systematic
708 review and interactive web app-based meta-analysis. *Ageing Res Rev*. 2023;90:102012.
- 709 [12] Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, et al. The
710 PRISMA extension statement for reporting of systematic reviews incorporating network
711 meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med*.
712 2015;162:777-84.
- 713 [13] Clark HD, Wells GA, Huët C, McAlister FA, Salmi LR, Fergusson D, et al. Assessing the
714 quality of randomized trials: reliability of the Jadad scale. *Control Clin Trials*.
715 1999;20:448-52.
- 716 [14] Olivo SA, Macedo LG, Gadotti IC, Fuentes J, Stanton T, Magee DJ. Scales to assess
717 the quality of randomized controlled trials: a systematic review. *Phys Ther*. 2008;88:156-
718 75.
- 719 [15] Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al.
720 Assessing the quality of reports of randomized clinical trials: is blinding necessary?
721 *Control Clin Trials*. 1996;17:1-12.
- 722 [16] Higgins JP, Savović J, Page MJ, Elbers RG, Sterne JA. Assessing risk of bias in a
723 randomized trial. *Cochrane handbook for systematic reviews of interventions*. 2019:205-
724 28.
- 725 [17] Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a
726 revised tool for assessing risk of bias in randomised trials. *Bmj*. 2019;366:l4898.
- 727 [18] Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations
728 for examining and interpreting funnel plot asymmetry in meta-analyses of randomised
729 controlled trials. *Bmj*. 2011;343:d4002.
- 730 [19] Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by
731 a simple, graphical test. *Bmj*. 1997;315:629-34.
- 732 [20] Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and
733 adjusting for publication bias in meta-analysis. *Biometrics*. 2000;56:455-63.
- 734 [21] Borenstein M, Hedges LV, Higgins JP, Rothstein HR. *Introduction to meta-analysis*:
735 John Wiley & Sons; 2021.
- 736 [22] Balduzzi S, Rücker G, Nikolakopoulou A, Papakonstantinou T, Salanti G, Efthimiou O,
737 et al. netmeta: An R package for network meta-analysis using frequentist methods.
738 *Journal of Statistical Software*. 2023;106:1-40.

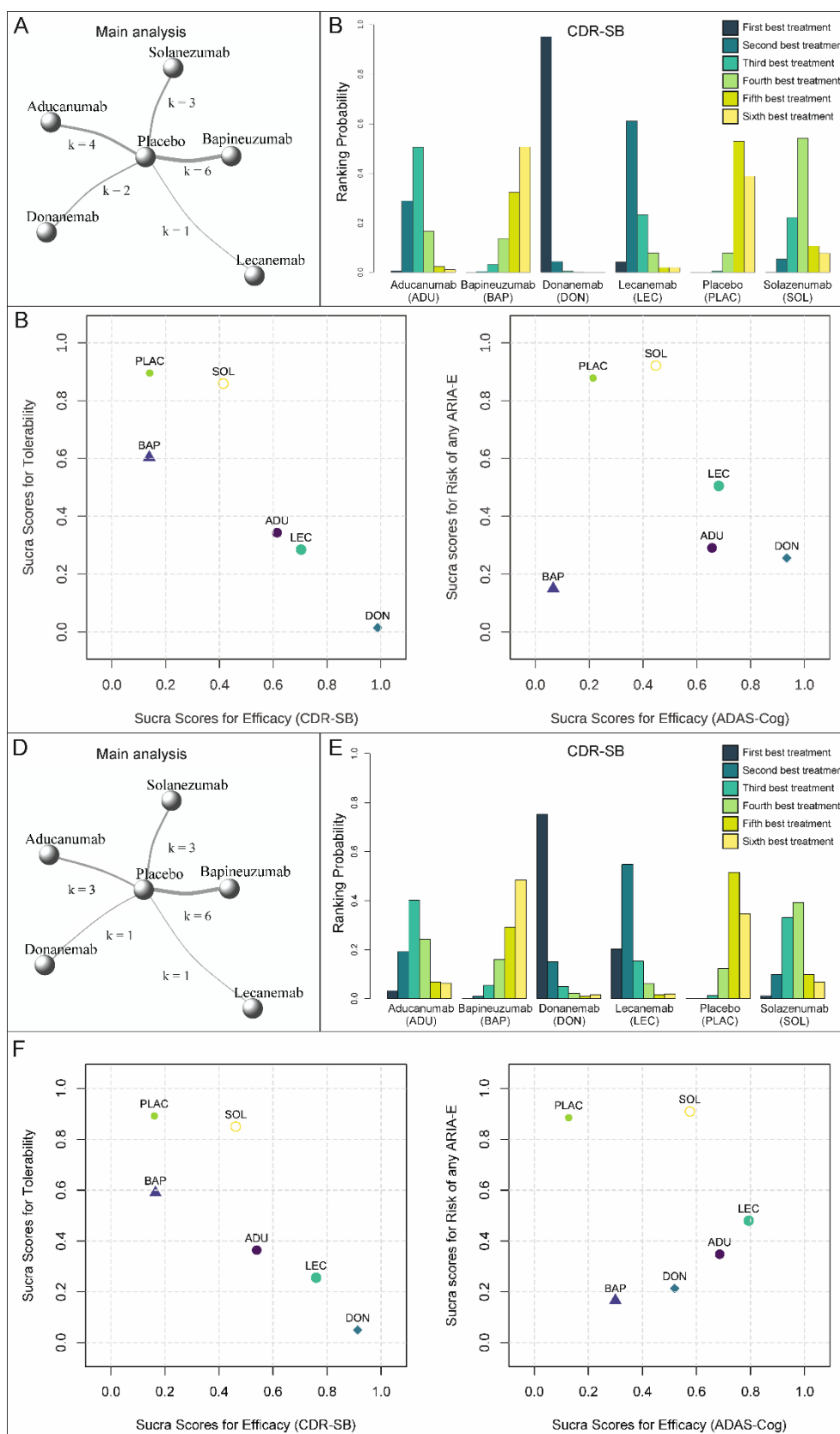
- 739 [23] Harrer M, Cuijpers P, Furukawa T, Ebert D. Doing meta-analysis with R: A hands-on
740 guide: Chapman and Hall/CRC; 2021.
- 741 [24] Rücker G, Schwarzer G. Ranking treatments in frequentist network meta-analysis
742 works without resampling methods. *BMC Med Res Methodol*. 2015;15:58.
- 743 [25] van Valkenhoef G, Lu G, de Brock B, Hillege H, Ades AE, Welton NJ. Automating
744 network meta-analysis. *Res Synth Methods*. 2012;3:285-99.
- 745 [26] Plummer M. JAGS: A program for analysis of Bayesian graphical models using Gibbs
746 sampling. Proceedings of the 3rd international workshop on distributed statistical
747 computing: Vienna, Austria; 2003. p. 1-10.
- 748 [27] Olkin I, Dahabreh IJ, Trikalinos TA. GOSH - a graphical display of study heterogeneity.
749 *Res Synth Methods*. 2012;3:214-23.
- 750 [28] Hartigan JA, Wong MA. Algorithm AS 136: A k-means clustering algorithm. *Journal*
751 *of the royal statistical society series c (applied statistics)*. 1979;28:100-8.
- 752 [29] Schubert E, Sander J, Ester M, Kriegel HP, Xu X. DBSCAN revisited, revisited: why and
753 how you should (still) use DBSCAN. *ACM Transactions on Database Systems (TODS)*.
754 2017;42:1-21.
- 755 [30] Fraley C, Raftery AE. Model-based clustering, discriminant analysis, and density
756 estimation. *Journal of the American statistical Association*. 2002;97:611-31.
- 757 [31] Cohen J. Weighted kappa: nominal scale agreement with provision for scaled
758 disagreement or partial credit. *Psychol Bull*. 1968;70:213-20.
- 759 [32] Daly CH, Neupane B, Beyene J, Thabane L, Straus SE, Hamid JS. Empirical evaluation
760 of SUCRA-based treatment ranks in network meta-analysis: quantifying robustness using
761 Cohen's kappa. *BMJ Open*. 2019;9:e024625.
- 762 [33] Foody GM. Explaining the unsuitability of the kappa coefficient in the assessment
763 and comparison of the accuracy of thematic maps obtained by image classification.
764 *Remote sensing of environment*. 2020;239:111630.
- 765 [34] Revelle W. An introduction to the psych package: Part I: data entry and data
766 description. Northwestern University. 2019.
- 767 [35] Salanti G. Indirect and mixed-treatment comparison, network, or multiple-
768 treatments meta-analysis: many names, many benefits, many concerns for the next
769 generation evidence synthesis tool. *Research synthesis methods*. 2012;3:80-97.
- 770 [36] Cipriani G, Ulivi M, Danti S, Lucetti C, Nuti A. Sexual disinhibition and dementia.
771 *Psychogeriatrics*. 2016;16:145-53.
- 772 [37] Cipriani A, Barbui C, Salanti G, Rendell J, Brown R, Stockton S, et al. Comparative
773 efficacy and acceptability of antimanic drugs in acute mania: a multiple-treatments
774 meta-analysis. *Lancet*. 2011;378:1306-15.
- 775 [38] Wu Y-C, Shih M-C, Tu Y-K. Using normalized entropy to measure uncertainty of
776 rankings for network meta-analyses. *Medical Decision Making*. 2021;41:706-13.
- 777 [39] Kraemer HC, Kupfer DJ. Size of treatment effects and their importance to clinical
778 research and practice. *Biol Psychiatry*. 2006;59:990-6.
- 779 [40] Salloway S, Sperling R, Fox NC, Blennow K, Klunk W, Raskind M, et al. Two phase 3
780 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. *N Engl J Med*.
781 2014;370:322-33.
- 782 [41] Vandenberghe R, Rinne JO, Boada M, Katayama S, Scheltens P, Vellas B, et al.
783 Bapineuzumab for mild to moderate Alzheimer's disease in two global, randomized,
784 phase 3 trials. *Alzheimers Res Ther*. 2016;8:18.

- 785 [42] Doody RS, Thomas RG, Farlow M, Iwatsubo T, Vellas B, Joffe S, et al. Phase 3 trials of
786 solanezumab for mild-to-moderate Alzheimer's disease. *N Engl J Med*. 2014;370:311-21.
- 787 [43] Honig LS, Vellas B, Woodward M, Boada M, Bullock R, Borrie M, et al. Trial of
788 Solanezumab for Mild Dementia Due to Alzheimer's Disease. *N Engl J Med*.
789 2018;378:321-30.
- 790 [44] Budd Haeberlein S, Aisen PS, Barkhof F, Chalkias S, Chen T, Cohen S, et al. Two
791 Randomized Phase 3 Studies of Aducanumab in Early Alzheimer's Disease. *J Prev*
792 *Alzheimers Dis*. 2022;9:197-210.
- 793 [45] van Dyck CH, Swanson CJ, Aisen P, Bateman RJ, Chen C, Gee M, et al. Lecanemab in
794 Early Alzheimer's Disease. *N Engl J Med*. 2023;388:9-21.
- 795 [46] Jeremic D, Navarro-López JD, Jiménez-Díaz L. Efficacy and Safety of Anti-Amyloid- β
796 Monoclonal Antibodies in Current Alzheimer's Disease Phase III Clinical Trials: A
797 Systematic Review and Interactive Web App-based Meta-Analysis. *Ageing Research*
798 *Reviews*. 2023:102012.
- 799 [47] Largent EA, Peterson A, Karlawish J, Lynch HF. Aspiring to Reasonableness in
800 Accelerated Approval: Anticipating and Avoiding the Next Aducanumab. *Drugs Aging*.
801 2022;39:389-400.
- 802 [48] Beshir SA, Aadithsoorya AM, Parveen A, Goh SSL, Hussain N, Menon VB.
803 Aducanumab Therapy to Treat Alzheimer's Disease: A Narrative Review. *Int J Alzheimers*
804 *Dis*. 2022;2022:9343514.
- 805 [49] Kuller LH, Lopez OL. ENGAGE and EMERGE: Truth and consequences? *Alzheimers*
806 *Dement*. 2021;17:692-5.
- 807 [50] Tampi RR, Forester BP, Agronin M. Aducanumab: evidence from clinical trial data
808 and controversies. *Drugs Context*. 2021;10.
- 809 [51] Shcherbinin S, Evans CD, Lu M, Andersen SW, Pontecorvo MJ, Willis BA, et al.
810 Association of Amyloid Reduction After Donanemab Treatment With Tau Pathology and
811 Clinical Outcomes: The TRAILBLAZER-ALZ Randomized Clinical Trial. *JAMA Neurol*.
812 2022;79:1015-24.
- 813 [52] Ross EL, Weinberg MS, Arnold SE. Cost-effectiveness of Aducanumab and
814 Donanemab for Early Alzheimer Disease in the US. *JAMA Neurol*. 2022;79:478-87.
- 815 [53] Touns K, Hathaway A, Gordon D, Chung H, Raji C, Boyd A, et al. Precision Medicine
816 Approach to Alzheimer's Disease: Successful Pilot Project. *J Alzheimers Dis*.
817 2022;88:1411-21.
- 818 [54] Cogswell PM, Barakos JA, Barkhof F, Benzinger TS, Jack CR, Jr., Poussaint TY, et al.
819 Amyloid-Related Imaging Abnormalities with Emerging Alzheimer Disease Therapeutics:
820 Detection and Reporting Recommendations for Clinical Practice. *AJNR Am J Neuroradiol*.
821 2022;43:E19-e35.
- 822 [55] Doran SJ, Sawyer RP. Risk factors in developing amyloid related imaging
823 abnormalities (ARIA) and clinical implications. *Front Neurosci*. 2024;18:1326784.
- 824 [56] Higgins JP. *Cochrane handbook for systematic reviews of interventions version 6.2*.
825 *Cochrane Database Syst Rev*. 2021.

826

827

828 **FIGURES**



829

830 **Figure 1.** The results of Bayesian NMA for efficacy and safety of anti-amyloid
 831 antibodies in Alzheimer's disease in phase III randomized placebo-controlled clinical
 832 trials. **A**, **B** and **C** correspond to the main analysis with all studies included, and **D**, **E**

833 and **F** refer to the sensitivity analysis with two influential studies left out: Donanemab
834 study in low/medium-*tau* population and high-dose EMERGE study of Aducanumab. **A**,
835 **D**. Network models with nodes (circles) representing antibodies, and the width of the
836 edges or links (*k*) indicating the number of studies included for each comparison. **B**, **E**.
837 Rankogram for global (cognitive and functional) outcome CDR-SB. The Y axis
838 represents the probability for each treatment being the best option, second best option,
839 third best option etc. (legend). **C**, **F** Scatterplots showing jointly the ranking results for
840 efficacy and safety of drugs in the network. SUCRA scores were calculated for the main
841 efficacy outcomes (X-axes: CDR-SB and ADAS-Cog) and safety measures (Y-axes:
842 tolerability and risk of any ARIA-E). Abbreviations: ARIA-E, Amyloid-Related
843 Imaging Abnormalities in form of cerebral Edema; CDR-SB, Clinical Dementia Rating
844 Sum of Boxes; SUCRA, Surface Under Cumulative RAnking curve.

845

846 **TABLES**

847

848 **Table 1.** League tables representing the results of random-effect model Frequentist and
 849 Bayesian network meta-analysis (NMA). The values in the tables represent
 850 Standardized Mean Differences (SMDs) for cognitive outcome Alzheimer's Disease
 851 Assessment Scale-Cognitive sub-scale (ADAS-Cog, **A, B**) and for global (cognitive and
 852 functional) outcome Clinical Dementia Rating Sum of Boxes (CDR-SB, **C, D**). The
 853 number below 0 suggests that the drug in the column is superior to the drug in the row.
 854 Darker green colors correspond to relatively greater effects, and bolded numbers denote
 855 statistical significance (**A, C**) or strong evidence (**B, D**). The numbers in parentheses
 856 represent confidence intervals (CIs) in Frequentist NMA (**A, C**) or credible intervals
 857 (CrIs) in Bayesian framework (**B, D**). Dagger symbol (†) denotes that the interpretation
 858 differs considerably between frequentist and Bayesian NMA results, indicating the
 859 uncertainty and lack of strong evidence for a non-zero effect.

860

Table 1A. Frequentist Network Meta-Analysis: ADAS-Cog

Aducanuma b					
-0.0960 (-0.1949; 0.003)	Bapineuzu mab				
0.0913 (-0.0444; 0.227)	0.1873 (0.0624; 0.3121)	Donanemab			
-0.0017 (-0.1251; 0.1217)	0.0943 (-0.0171; 0.2057)	-0.0930 (-0.238; 0.052)	Lecanemab		
-0.1174 * (-0.1968; - 0.038)	-0.0214 (-0.0805; 0.0376)	-0.2087 ** (-0.3187; - 0.0987)	-0.1157 † *** (-0.2101; - 0.0213)	Placebo	
-0.0503 (-0.1502; 0.0497)	0.0457 (-0.039; 0.1304)	-0.1416 † (-0.2672; - 0.0159)	-0.0486 (-0.1609; 0.0637)	0.0671 † **** (0.0064; 0.1278)	Solanezu mab

Statistical significance for Drug vs. Placebo: * p = 0.0038; ** p = 0.0002; *** p = 0.0163; **** p = 0.0303

Heterogeneity I² = 0% [0.0%; 58.3%]; p-value of Q statistics = 0.6985 in heterogeneity (within designs).

Table 1B. Bayesian Network Meta-Analysis: ADAS-Cog

Aducanuma b	
-0.0959	Bapineuzu mab

(-0.2065; 0.0178)					
0.0853 (-0.068; 0.2344)	0.1803 (0.0386; 0.3196)	Donanemab			
-0.0025 (-0.1519; 0.1465)	0.0932 (-0.0461; 0.2342)	-0.0870 (-0.2589; 0.0872)	Lecanemab		
-0.1176 (-0.2058; -0.0288)	-0.0218 (-0.0890; 0.0446)	-0.2026 (-0.3246; -0.0778)	-0.1154 † (-0.2371; 0.0060)	Placebo	
-0.04958 (-0.16684; 0.0646)	0.0454 (-0.0569; 0.1468)	-0.1358 † (-0.2784; 0.0108)	-0.0479 (-0.1905; 0.0947)	0.0676 † (-0.0092; 0.1421)	Solanezu mab

Table 1C. Frequentist Network Meta-Analysis: CDR-SB

Aducanumab					
-0.1053 † (-0.2073; -0.0033)	Bapineuzumab				
0.2046 (0.0622; 0.3471)	0.3099 (0.1791; 0.4407)	Donanemab			
0.0357 (-0.0977; 0.1690)	0.1409 † (0.0202; 0.2617)	-0.1690 † (-0.3255; -0.0125)	Lecanemab		
-0.1004 * (-0.1829; -0.0180)	0.0048 (-0.0552; 0.0649)	-0.3051 ** (-0.4213; -0.1889)	-0.1361 *** (-0.2409; -0.0313)	Placebo	
-0.0473 (-0.1547; 0.0602)	0.0580 (-0.0334; 0.1494)	-0.2519 (-0.3870; -0.1168)	-0.0829 (-0.2083; 0.0425)	0.0532 (-0.0157; 0.1220)	Solanezu mab

Statistical significance: Drug vs. Placebo * p = 0.017; ** p < 0.0001; *** p = 0.0109

Heterogeneity I² = 9.8% [0.0%; 48.8%] p-value of Q statistics = 0.3493 in heterogeneity (within designs).

Table 1D. Bayesian Network Meta-Analysis: CDR-SB

Aducanumab			
-0.1045 † (-0.2215, 0.0142)	Bapineuzumab		
0.2022 (0.0361, 0.3628)	0.3063 (0.1493, 0.4552)	Donanemab	
0.0351 (-0.1328, 0.2049)	0.1397 † (-0.0191, 0.2954)	-0.1672 † (-0.3579, 0.033)	Lecanemab

-0.101 (-0.1961, - 0.0073)	0.0037 (-0.0683, 0.0736)	-0.3026 (-0.4359, - 0.1642)	-0.1358 (-0.2771, 0.0032)	Placebo	
-0.0499 (-0.1814, 0.0762)	0.0547 (-0.0615, 0.164)	-0.2522 (-0.4118, - 0.0918)	-0.0849 (-0.2558, 0.0779)	0.0511 (-0.0387, 0.1372)	Solanezu mab

861

862

863

864 **Table 2.** League tables representing the results of random-effect model Frequentist
 865 NMA and Bayesian NMA. The values in the tables represent Relative Risks (natural
 866 logarithm) for tolerability (treatment discontinuations due to adverse events, **A, B**) and
 867 for amyloid-related imaging abnormalities in form of cerebral edema (ARIA-E, **C, D**).
 868 For each comparison, the number below 0 suggests that the drug in the column is safer
 869 than the drug in the row. Darker red colors correspond to relatively higher risk. Bolded
 870 numbers indicate statistical significance for Frequentist NMA (**A, C**), or strong evidence
 871 in Bayesian NMA (**B, D**). The numbers in parentheses represent confidence intervals
 872 (CIs) for Frequentist NMA (**A, C**), or credible intervals (CrIs) in Bayesian framework
 873 (**B, D**). Dagger symbol (†) denotes when the interpretation differs significantly between
 874 frequentist and Bayesian NMA results, indicating the uncertainty and lack of strong
 875 evidence for a non-zero risk.

876

877

Table 2A. Frequentist Network Meta-Analysis: Tolerability (Treatment Discontinuations due to Adverse Events)

Aducanuma b					
0.2939 (-0.0956; 0.6835)	Bapineuzu mab				
-0.6066 (-1.0984; - 0.1148)	-0.9007 (-1.3042; - 0.4969)	Donanema b			
-0.1013 (-0.6663; 0.4639)	-0.3951 (-0.8855; 0.095)	0.5054 (-0.0695; 1.0802)	Lecanema b		
0.4973 * (0.1576; 0.837)	0.2033 † ** (0.0126; 0.3941)	1.1039 *** (0.7483; 1.4596)	0.5985 **** (0.1469; 1.0501)	Placebo	
0.4879 (0.0683; 0.9075)	0.194 (-0.1175; 0.5055)	1.0945 (0.6619; 1.5272)	0.5891 (0.0747; 1.1036)	-0.0093 (-0.2557; 0.237)	Solanezu mab

Statistical significance: Drug vs. Placebo: * p = 0.0041; ** p = 0.0367; *** p < 0.0001; **** p = 0.0094

Heterogeneity I² = 0% [0.0%; 58.3%]; p-value of Q statistics = 0.9681 in heterogeneity (within designs).

Table 2B. Bayesian Network Meta-Analysis: Tolerability (Treatment Discontinuations due to Adverse Events)

Aducanuma b	
0.3021	Bapineuzu mab

(-0.1091, 0.7240)					
-0.6015 (-1.1231, -0.0708)	-0.9037 (-1.3593, -0.4631)	Donanema b			
-0.0962 (-0.7213, 0.5193)	-0.4024 (-0.9524, 0.1380)	0.5020 (-0.1335, 1.1416)	Lecanemab		
0.4985 (0.1378, 0.8585)	0.1973 † (-0.0195, 0.4092)	1.1006 (0.7128, 1.5001)	0.5993 (0.0926, 1.1026)	Placebo	
0.4882 (0.0415, 0.9542)	0.1879 (-0.1697, 0.5445)	1.0923 (0.6117, 1.5767)	0.5932 (0.0163, 1.1629)	-0.0093 (-0.2870, 0.2734)	Solanezu mab

Table 2C. Frequentist Network Meta-Analysis: Total events of cerebral edema (ARIA-E)

Aducanuma b					
-0.1948 (-0.862; 0.4723)	Bapineuzu mab				
-0.0471 (-0.7302; 0.6361)	0.1477 (-0.6694; 0.9648)	Donanema b			
0.4015 (-0.3838; 1.1866)	0.5963 (-0.3077; 1.5004)	0.4486 (-0.4674; 1.3646)	Lecanemab		
2.4197 * (2.0703; 2.7691)	2.6145 * (2.0462; 3.1828)	2.4668 * (1.8797; 3.0539)	2.0182 * (1.3151; 2.7213)	Placebo	
2.5416 (1.7171; 3.3662)	2.7365 (1.798; 3.6749)	2.5887 (1.6388; 3.5387)	2.1402 (1.1145; 3.1659)	0.122 (-0.6249; 0.8688)	Solanezu mab

Statistical significance for Drug vs. Placebo: * $p < 0.0001$.

Heterogeneity $I^2 = 25.6\%$ [0.0%; 63.1%]; p-value of Q statistics = 0.1997 in heterogeneity (within designs).

Table 2D. Bayesian Network Meta-Analysis: Total events of cerebral edema (ARIA-E)

Aducanuma b		
-0.2148 (-1.1414, 0.5154)	Bapineuzu mab	
-0.0492	0.1619	Donanema b

(-0.9452, 0.8161)	(-0.7384, 1.2530)				
0.4018 (-0.6774, 1.4623)	0.6111 (-0.4272, 1.9019)	0.4492 (-0.7444, 1.6828)	Lecanemab		
2.4186 (1.9414, 2.8973)	2.6376 (2.0373, 3.3955)	2.4691 (1.7415, 3.2262)	2.0176 (1.0681, 2.9953)	Placebo	
2.5342 (1.4692, 3.4767)	2.7477 (1.6868, 3.8665)	2.5816 (1.3860, 3.6877)	2.1355 (0.7887, 3.3623)	0.1090 (-0.8196, 0.9406)	Solanezu mab

