Supplemental Online Content

Hu Y, Kirmess KM, Meyer MR, et al. Assessment of a plasma amyloid probability score to estimate amyloid positron emission tomography findings among adults with cognitive impairment. *JAMA Netw Open*. 2022;5(4):e228392. doi:10.1001/jamanetworkopen.2022.8392

This supplemental material has been provided by the authors to give readers additional information about their work.

© 2022 Hu Y et al. *JAMA Network Open*.

eMethods.

1.1 Amyloid PET image analyses

In PARIS-DISCOVERY, amyloid PET tracer uptake (SUVr) was quantified using automated scan processing procedures that did not require reviewer input and included the following brain regions: gyrus rectus, prefrontal, anterior cingulate, parietal, posterior cingulate, precuneus cortices, and caudate nucleus head. Using 510(k)–cleared and CE-marked medical imaging workstations, multi-planar images were displayed, co-registered, assessed and analyzed by each reviewer using the automated workflow that generated SUVr values (using whole cerebellum as reference region) from the Florbetapir Clark atlas regions and automatically converted these values into the Centiloid scale. All standardized images, published SUVr reference data, and documentation used to calculate Centiloid values for the three tracers can be found on the GAAIN website (http://www.gaain.org/centiloid-project). The following equations were used to convert the mean composite SUVr to the Centiloid scale for each tracer as derived by American College of Radiology: Florbetapir: 183 * mean composite SUVr - 177; Florbetaben: 153.4 * mean composite SUVr - 154.9; Flutemetamol: 121.42 * mean composite SUVr - 121.16.

In the MissionAD data set, two amyloid PET tracers were used: Florbetapir and Florbetaben. Amyloid PET SUVr values were calculated as mean composite SUVr (a simple average of cingulate, frontal, parietal, and temporal cortices) using whole cerebellum as reference region. The following equations were used to convert the mean composite SUVr to the Centiloid scale for each tracer as derived by Bioclinica (Adamczuk et al., 2019; Klunk et al., 2015): Florbetapir: 205.72 * mean composite SUVr - 209.63; Florbetaben: 175.57 * mean composite SUVr - 173.21.In the MissionAD data set analyzed, two amyloid PET tracers were used: Florbetapir and Florbetaben. Amyloid PET SUVr values were calculated as mean composite SUVr (a simple average of cingulate, frontal, parietal, and temporal cortices) using whole cerebellum as reference region. The following equations were used to convert the mean composite SUVr to the Centiloid scale for each tracer as derived by Bioclinica (Adamczuk et al., 2019; Klunk et al., 2015): Florbetapir: 205.72 * mean composite SUVr - 209.63; Florbetaben: 175.57 * mean composite SUVr - 173.21.

1.2 Plasma Aβ42/40 ratio determination

Briefly, USP-traceable amino acid analysis was conducted on full length ^{14}N- and uniformly labeled ^{15}N- Aβ40 and ^{15}N-Aβ42 proteins (rPeptide, Watkinsville, GA) to confirm their chemical purity and amount to be used when preparing calibrator stock concentrations.

© 2022 Hu Y et al. *JAMA Network Open*.

Calibrator concentrations were prepared based on the expected physiological range for plasma Aβ proteins: Aβ40 = 24.3 – 1558 pg/mL; Aβ42 = 3.6 – 235 pg/mL.

On the day of sample analysis, participant plasma, quality control plasma samples, frozen calibrators, and uniformly labeled ^{15}N full length Aβ40 and Aβ42 internal standards were thawed. All calibrators, quality control samples, and participant samples were treated identically throughout sample processing and analysis. To each 450 µL of sample, an immunoprecipitation buffer containing ^{15}N-Aβ40 and ^{15}N-Aβ42 was added prior to immunocapture. Aβ was extracted from the samples by immunoprecipitation using a monoclonal antibody, HJ5.1 (Aβ amino acid residues 13-28). After immunocapture, the Aβ-bound to magnetic beads was digested using Lys-N metalloendoprotease (Thermo Fisher Scientific, Waltham, MA). The Lys-N digested Aβ species into Aβ peptides Aβ28-40 (amino acid sequence KGAIIGLMVGGVV) and Aβ28-42 (amino acid sequence KGAIIGLMVGGVVIA). The Aβ digests were further purified by reverse-phase solid phase extraction (SPE) (Waters Corp., Milford, MA). These eluted Aβ peptides were concentrated and dried under vacuum before being reconstituted in 10% ACN/10% formic acid prior to injection onto the LC-MS/MS system.

The 96-well collection plate containing the resolubilized Aβ peptides was placed in the temperature-controlled LC autosampler (Waters Acquity LC). Three µL from each well were injected onto a monolithic divinylbenzene column (Thermo Fisher Scientific) where Aβ28-40 and Aβ28-42 were separated, identified, and quantified using LC-MS/MS (Acquity UPLC M-Class liquid chromatography unit (Waters Corp.) interfaced to a Thermo Fisher Scientific Fusion Lumos Mass Spectrometer).

The total peak area for the endogenous ^{14}N Aβ peptides were divided by the total peak area for the exogenously added, uniformly labeled ^{15}N Aβ peptide internal standards to obtain a peak area ratio. The peak area ratio for each Aβ peptide indicated the peptide concentration based on an external eight-point standard curve that spanned the expected physiological range, and Aβ peptide concentrations (pg/mL) were determined from the standard curve. The measured concentrations were expressed as the plasma Aβ42/40 ratio. The LC-MS/MS data were assembled and assessed using TraceFinder 4.1 General Quan software (Thermo Fisher Scientific).

1.3 APOE proteotype determination

Briefly, samples were diluted 1:20 in diluent buffer containing denaturing and reducing agents (95 µL 100 mM Tris pH 8.1, 9.6 mM SDC, 2.3 mM TCEP). Samples were further diluted 1:6 in the same diluent spiked with stable isotope labeled internal standard

© 2022 Hu Y et al. *JAMA Network Open*.

peptides. The samples were heat denatured (50ºC), then alkylated with iodoacetamide. After these successive steps, the samples were digested for 90 min at 50 °C. The samples were acidified to terminate digestion and precipitate the denaturing agent. The digested samples were cleaned by solid phase extraction, dried in a SpeedVac vacuum concentrator (Thermo Fisher Scientific), and reconstituted in LC-MS buffer for analysis.

The digested APOE peptide samples were separated and analyzed on the same LC-MS/MS instrumentation as the Aβ samples. Three μL of sample were injected onto a CSH C18 column (Waters Corp.). Following chromatographic separation, the APOE peptides were introduced into the MS inlet via electrospray ionization, where precursor ions were filtered using a quadrupole isolation window of 1.6 m/z. Product ions were detected within the orbitrap at a mass resolution of 30,000 and the AGC target was set to $5.0e^5$. The product ions were detected, and the peak areas of each fragment ion were determined post-acquisition using Skyline software (MacCoss Lab, University of Washington, Seattle, WA).

APOE proteotyping for each of the six *APOE* genotypes (APOE2/2, APOE2/3, APOE2/4, APOE3/3, APOE3/4, APOE4/4) used a combination of the presence or absence of the four targeted APOE isoform-specific tryptic peptides. The presence or absence of an APOE isoform-specific peptide was determined by monitoring the peak areas for each peptide. An R-script was utilized to generate the APOE proteotypes based upon the input peak areas (or their absence) for each isoform-specific peptide (The R Foundation for Statistical Computing, https://www.r-project.org/).

1.4 Statistical analysis, algorithm development, and algorithm performance testing

All data analysis was performed using R version 4.0.0 (The R Foundation for Statistical Computing). Receiver operating characteristic (ROC) analyses, including the calculation of the area under the ROC curve (AUC), sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were conducted using the pROC package for R and optimal cutoff values for plasma Aβ42/40 ratio and model parameters were determined by the Youden index (maximized sensitivity and specificity of the predictive test) (Robin et al., 2011). 95% confidence intervals (CI) for AUC and comparisons between ROCs were calculated using the DeLong method (DeLong et al., 1988). Comparisons between groups with only two outcomes were performed using an unpaired two-sided t-test. Confusion matrices and associated statistics (specifically accuracy and accuracy confidence intervals) were generated using the confusion matrix function from the caret package (Kuhn, 2008).

Binary logistic regression models (LRM) were used to predict amyloid positivity

© 2022 Hu Y et al. *JAMA Network Open*.

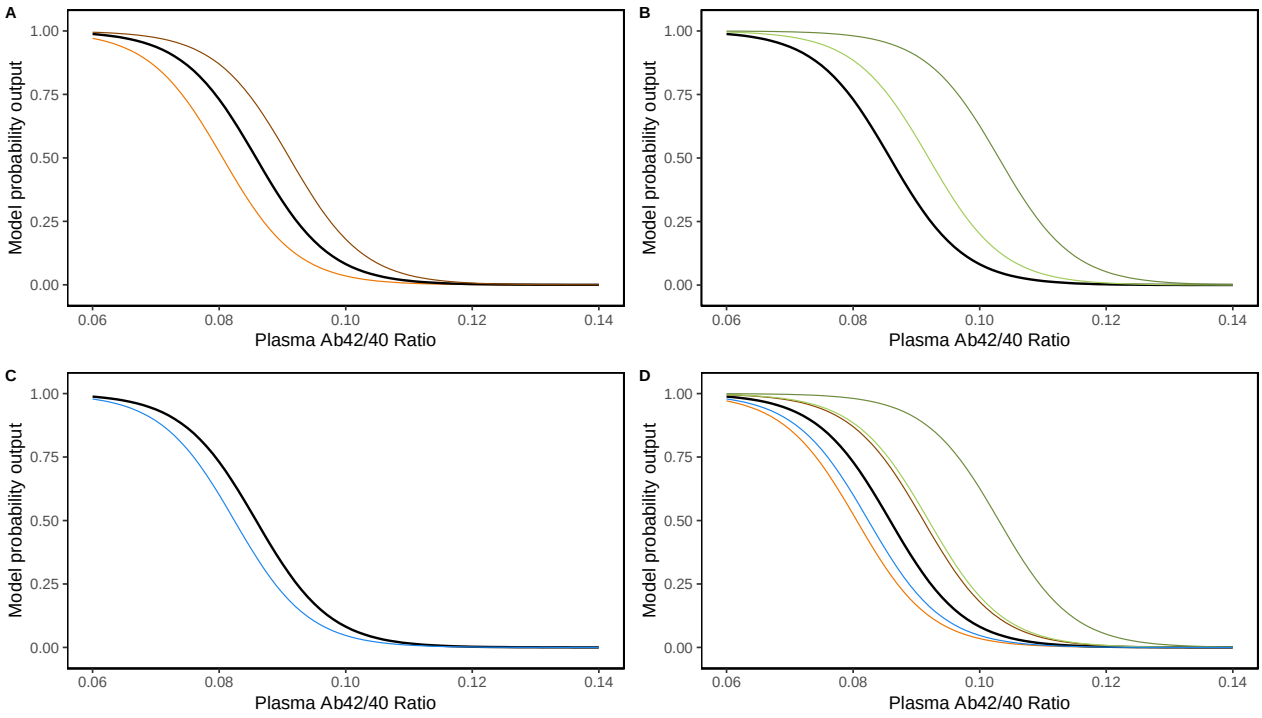
based on plasma Aβ42/40 ratio, APOE proteotype, and age, using amyloid positivity (defined as a Centiloid score > 25) as the dependent variable and plasma Aβ42/40 ratio, APOE proteotype, and age as independent variables. APOE proteotype was included as three dummy variables indicating the presence of either one or two APOE4 alleles (two dummy variables) and presence of any APOE2 alleles (one dummy variable). This method allows for a non-linear influence of two vs. one APOE4 allele.

Validation of the model developed on PARIS-DISCOVERY in the MissionAD cohort was performed by using the PARIS-DISCOVERY LRM to calculate probabilities for all samples in the MissionAD cohort. A calibration curve was generated from these probabilities using the val.prob.ci.2 function in the CalibrationCurves package (Van Calster et al 2016).

Van Calster B, Nieboer D, Vergouwe Y, De Cock B, Pencina MJ, Steyerberg EW. A calibration hierarchy for risk models was defined: from utopia to empirical data. J Clin Epidemiol. 2016 Jun;74:167-76. doi: 10.1016/j.jclinepi.2015.12.005. Epub 2016 Jan 6. PMID: 26772608.

© 2022 Hu Y et al. *JAMA Network Open*.

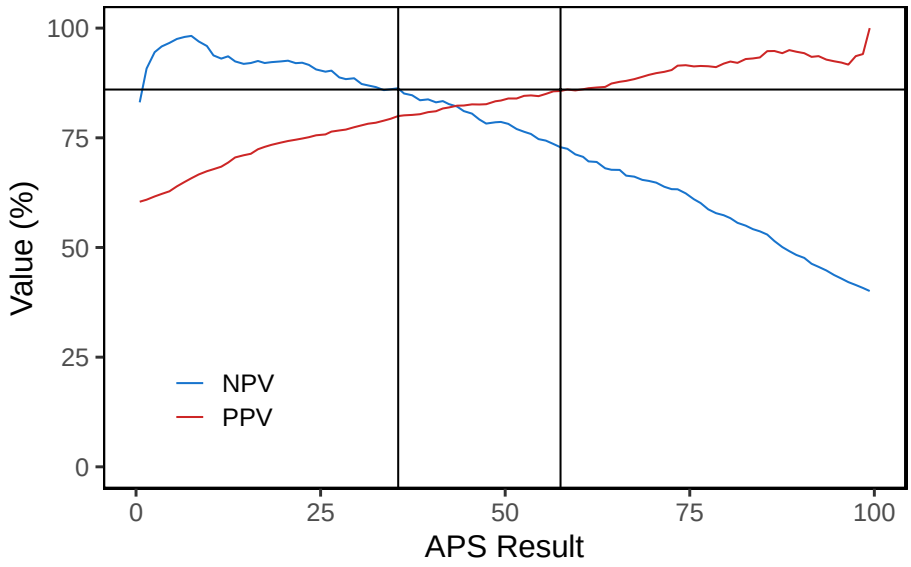
eFigure 1. Effects of Inputs into Model



The model trained on the combined dataset was used to generate probability outputs for all Aβ42/40 ratios between 0.06 and 0.14. The other model inputs for this base case patient were set to age = 70 and APOE genotype = E3/E3 (black line on all plots). The black circles at the top of the plots mark Aβ42/40 ratios for amyloid-positive patients and the points at the bottom the Aβ42/40 ratios for amyloid-negative patients, to avoid overlap and illustrate density the y-axis position of the points has a variable component. The effect of increasing or decreasing age by 10 years is shown in panel A (orange lines is for base case patient with age set to 60, brown line is for age 80). The effect of one or two APOE4 alleles is shown in panel B (light green line is for base case patient with genotype of E3/E4, and darker green line for genotype of E4/E4). This illustrates that having two copies of E4 has a greater effect on the calculated probability than having only one copy of E4. Panel C shows the effect of having an APOE2 allele (light blue line is base case patient but with E3/E2 genotype). When all the lines are combined, it is possible to compare the effects in the model that are attributed to age and APOE genotype. In general, one APOE4 allele has a similar effect to being slightly more than 10 years older (compare light green line to brown line). And the presence of the APOE2 allele is like being a little less than 10 years younger (compare blue line to orange line).

© 2022 Hu Y et al. *JAMA Network Open*.

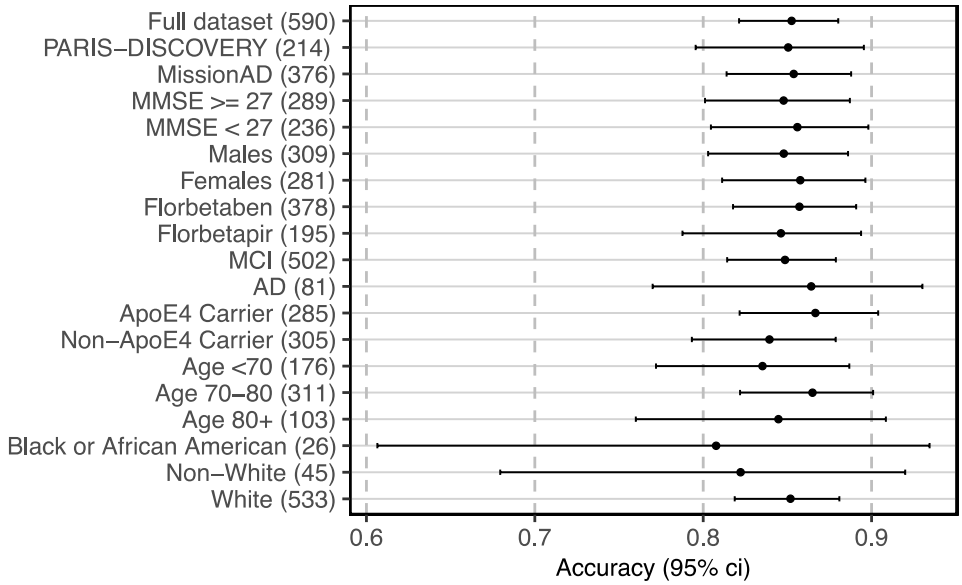
eFigure 2. NPV and PPV at Various APS Values as Calculated for a Population With 60% Prevalence of PET Amyloid Positivity



Horizontal line is at 86% (the target NPV and PPV), and vertical lines illustrate the 35 and 58 APS cut points.

© 2022 Hu Y et al. *JAMA Network Open*.

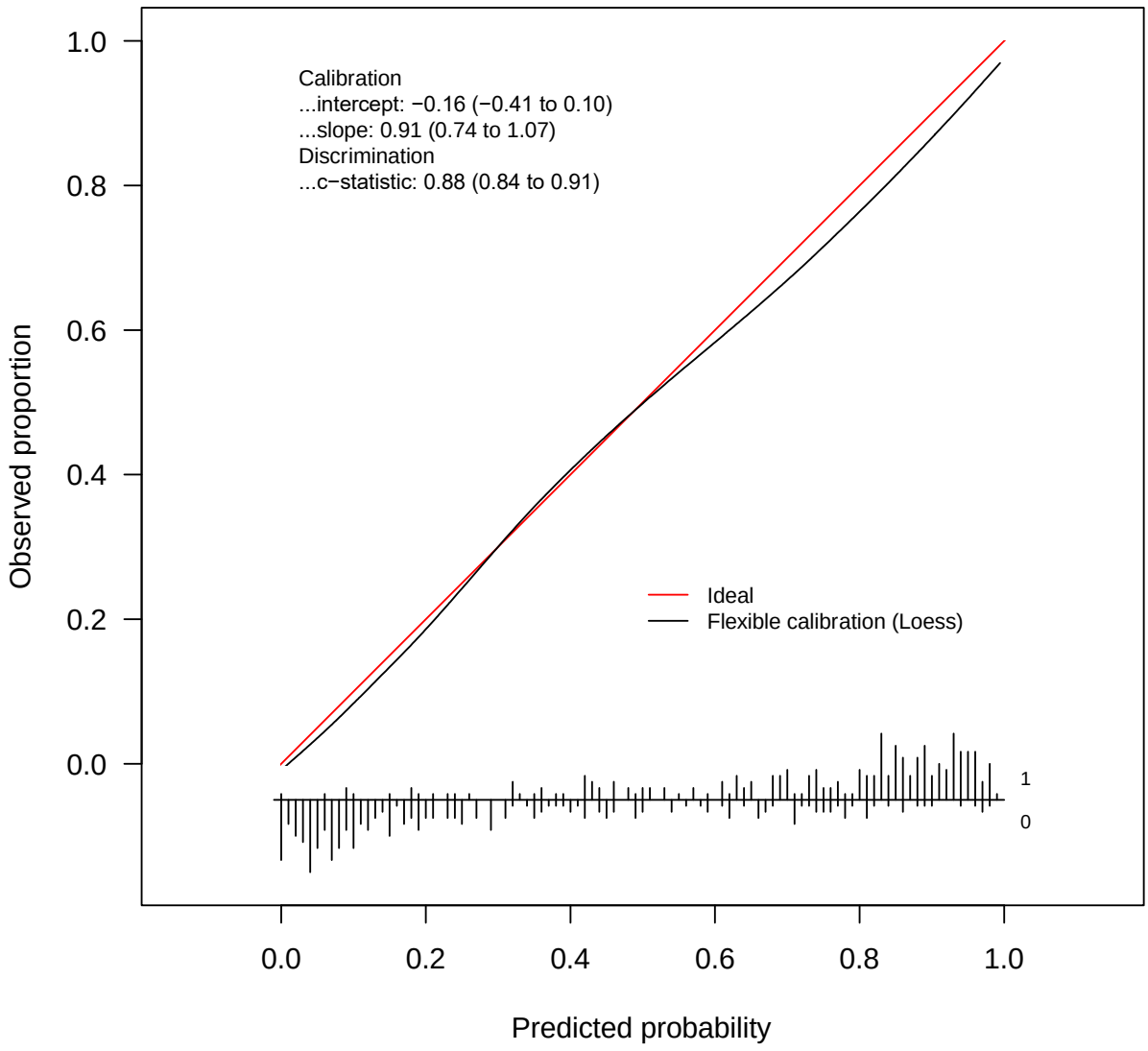
eFigure 3. Accuracy of APS in Subgroups of the Pooled Dataset



For accuracy comparison, the patients in the intermediate group are excluded and the APS classification is compared to the amyloid PET classification. The number in parentheses represents the number of patients in each subgroup, less those classified in the intermediate group. Accuracy was compared pairwise as follows by Fisher's exact test: Males/Females, Florbetaben/Florbetapir, and APOE4 carriers/non-carriers. These pairwise comparisons showed that there were no significant differences in the observed accuracy in these subgroups.

© 2022 Hu Y et al. *JAMA Network Open*.

eFigure 4. Calibration of Model Developed on PARIS Discovery Cohort When Applied to MissionAD Cohort



The LRM developed on PARIS-DISCOVERY was used to generate predicted probabilities on the samples from the MissionAD cohort. This calibration plot shows the observed proportion (prevalence) vs. the PARIS-DISCOVERY trained model predictions. The slope for this calibration curve is close to 1 (0.91) and the intercept is close to zero (-0.16). This demonstrates that the predicted probabilities agree well with the observed proportions and that the model developed on PARIS-DISCOVERY performs well in the MissionAD cohort.

© 2022 Hu Y et al. *JAMA Network Open*.

eTable 1. Eligibility Criteria for PARIS Discovery and MissionAD

	PARIS-DISCOVERY	**MissionAD**
Recruitment period	Phase1: from 22OCT2018 to 05JAN2019	From 3rd quarter 2016 through 1st quarter 2019
Eligible population	Clinical diagnosis of MCI or dementia (without prior confirmation by amyloid PET)	Clinical diagnosis of MCI due to AD or mild AD (regardless of prior confirmation by amyloid PET)
Age eligibility	≥65 years	50 to 85 years
Diagnostic Criteria for MCI and mild AD	• DSM-IV and/or NIA-AA criteria, verified by a dementia specialist within 24 months; • Meets Appropriate Use Criteria: – Cognitive complaint with objectively confirmed impairment; – The etiologic cause of cognitive impairment is uncertain after a comprehensive evaluation by a dementia specialist, including general medical and neurological examination, mental status testing including standard measures of cognitive impairment, laboratory testing, and structural neuroimaging as below; – AD is a diagnostic consideration; – Knowledge of amyloid PET status is expected to alter diagnosis and management.	• Meets NIA-AA criteria plus the following: – MMSE score ≥24 – CDR global score of 0.5 – CDR Memory Box score 0.5 or greater • Cognitive impairment • A history of subjective memory decline with gradual onset and slow progression over the past year
Amyloid PET timing	Within the previous 18 months	Within 80 days of baseline screening (pre-randomization)
Amyloid PET reading	Blinded, centralized review conducted by ACR: two Readers plus one Adjudicator for visual read. Centiloid automatically derived by ACR's workflow.	Blinded, centralized review conducted by Bioclinica: one visual reader per scan; discussion with two other colleagues in difficult cases. Centiloid derived by Bioclinica workflow.

© 2022 Hu Y et al. *JAMA Network Open*.

eTable 2. Diagnostic Performance of APS in the Pooled Dataset

	Sensitivity	Specificity	Accuracy
Intermediate excluded	92%	77%	85%
Intermediate counted as positive	93%	65%	80%
Intermediate counted as negative	80%	80%	80%

eTable 3. Comparison of Odd Ratios for the Models Trained on PARIS Discovery, MissionAD, or the Combined Cohorts

	PARIS-DISCOVERY	MissionAD	Combined
0.010 lower Aβ42/40	5.56 (10.7, 2.88)	5.54 (8.2, 3.74)	5.52 (7.69, 3.96)
One copy of ApoE4	5.19 (2.44, 11.1)	2.07 (1.23, 3.48)	2.83 (1.87, 4.3)
Two copies of ApoE4	12.9 (2.74, 61.2)	31 (3.91, 246)	18.9 (5.51, 64.5)
Presence of ApoE2	1.05 (0.357, 3.07)	0.326 (0.13, 0.82)	0.556 (0.283, 1.09)
10-year older age	2.24 (1.23, 4.06)	2.57 (1.7, 3.9)	2.46 (1.76, 3.44)

For continuous variables the odds ratio is listed for a meaningful change in the variable. Table shows the odds ratio with 95% CI in parenthesis.

© 2022 Hu Y et al. *JAMA Network Open*.