Crowdsourced estimation of cognitive decline and resilience in Alzheimer's disease

AUTHORS: Genevera I Allen[1], Nicola Amoroso[2,3], Catalina Anghel[4], Venkat Balagurusamy[5], Christopher J Bare[6], Derek Beaton[7], Roberto Bellotti[2,3], David A Bennett[8], Kevin Boehme[9], Paul C Boutros[4, 53, 54], Laura Caberlotto[10], Cristian Caloian[4], Frederick Campbell[1], Elias Chaibub Neto[6], Yu-Chuan Chang[11], Beibei Chen[12], Chien-Yu Chen[13], Ting-Ying Chien[14], Tim Clark[15,16], Sudeshna Das[15,16], Christos Davatzikos[17], Jieyao Deng[18,19], Donna Dillenberger[5], Richard JB Dobson[20,21], Qilin Dong[18,19], Jimit Doshi[17], Denise Duma[22], Rosangela Errico[23], Guray Erus[17], Evan Everett[1], David W Fardo[24,25], Stephen H Friend[6], Holger Fröhlich[26], Jessica Gan[1], Peter St George-Hyslop[27], Satrajit S Ghosh[28,29], Enrico Glaab[30], Robert C Green[31], Yuanfang Guan[32,33,34], Ming-Yi Hong[13], Chao Huang[35], Jinseub Hwang[37], Joseph Ibrahim[35], Paolo Inglese[38], Qijia Jiang[1], Yuriko Katsumata[25], John SK Kauwe[9#], Arno Klein[6#], Dehan Kong[35], Roland Krause[30], Emilie Lalonde,[4] Mario Lauria[10], Eunjee Lee[35], Xihui Lin[4], Zhandong Liu[1], Julie Livingstone[4], Benjamin A Logsdon[6], Simon Lovestone[39], Anandhi Lyappan[40,26], Michelle Ma[12], Ashutosh Malhotra[40,26], Lara M Mangravite[6#], Taylor J Maxwell[41], Emily Merrill[15], John Nagorski[1], Aishwarya Namasivayam[30], Manjari Narayan[1], Mufassra Naz[40.26], Stephen J Newhouse[20,42], Thea C Norman[6], Ramil N Nurtdinov[43], Yen-Jen Oyang[11], Yudi Pawitan[36], Shengwen Peng[18,19], Mette A Peters[6#], Stephen R Piccolo[44], Paurush Praveen[10,26], Corrado Priami[10], Veronica Y Sabelnykova[4], Philipp Senger[40], Xia Shen[45], Andrew Simmons[20], Aristeidis Sotiras[17], Gustavo Stolovitzky[46,5], Sabina Tangaro[3], Andrea Tateo[2], Yi-An Tung[47], Nicholas J Tustison[48], Erdem Varol[17], George Vradenburg[49], Michael W Weiner[50], Guanghua Xiao[12], Lei Xie[51], Yang Xie[12], Jia Xu[12], Hojin Yang[35], Xiaowei Zhan[12], Yunyun Zhou[12], Fan Zhu[32], Hongtu Zhu,[35] Shanfeng Zhu[18,19,52], Alzheimer's Disease Neuroimaging Initiative*

*Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at:
http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Authors to whom correspondence should be addressed.

AFFILIATIONS:
[1]Department of Statistics and Electrical and Computer Engineering, Rice University, 6100 Main Street, Houston, TX, USA
[2]Dipartimento di Fisica "M. Merlin", Università degli studi di Bari "A. Moro", Via Orabona, 4, Bari, Italy
[3]Sezione di Bari, Istituto Nazionale di Fisica Nucleare, Via Orabona, 4, Bari, Italy
[4]Ontario Institute for Cancer Research, Informatics and Bio-computing Program, MaRS Centre, 661 University Avenue, Suite 510, Toronto, ON, Canada
[5]IBM Computational Biology Center, IBM Research, 1101 Kitchawan Road, Yorktown Heights, NY, USA
[6]Sage Bionetworks, 1100 Fairview Avenue North, M1-C108, Seattle, WA, USA
[7]School of Behavioral and Brain Sciences, The University of Texas at Dallas, 800 W. Campbell Road, Richardson, TX, USA
[8]Rush Alzheimer's Disease Center, Rush University Medical Center, 600 South Paulina, Suite 1028, Chicago, IL, USA
[9]Department of Biology, Brigham Young University, 4102 LSB, Provo, UT, USA
[10]The Microsoft Research- University of Trento Centre- COSBI, Piazza Manifattura 1 , Rovereto, Italy
[11]Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd. Taipei, Taiwan

[12]Quantitative Biomedical Research Center, The University of Texas Southwestern Medical Center, Suite [13]NC8.512 5323 Harry Hines Blvd. Dallas, TX, USA

Department of Bio-Industrial Mechatronics Engineering, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd. Taipei, Taiwan

[14]Innovation Center for Big Data and Digital Convergence, Yuan Ze University, 135 Yuan-Tung Road, Taoyuan, Taiwan

[15]Department of Neurology, Massachusetts General Hospital, 65 Landsdowne St, Suite 200, Cambridge, MA, USA

[16]Department of Neurology, Harvard Medical School, 25 Shattuck St, Boston, MA USA

[17]Center for Biomedical Image Computing and Analytics, University of Pennsylvania, 3600 Market Street, Philadelphia, PA, USA

[18]School of Computer Science , Fudan University, 220 Handan Road, Shanghai, Shanghai, China

[19]Shanghai Key Lab of Intelligent Information Processing, Fudan University, 220 Handan Road, Shanghai, Shanghai, China

[20]NIHR Biomedical Research Centre for Mental Health, Kings College London, Box P096, De Crespigny Park, London, UK

[21]Institute of Psychiatry, Psychology and Neuroscience, MRC Social, Genetic and Developmental Psychiatry Centre, Kings College London, Box P096, De Crespigny Park, London, UK

[22]Department of Pediatrics-Neurology, Baylor College of Medicine, 1250 Moursund street, Houston, TX, USA

[23]Università degli Studi di Genova, via Dodecaneso 33, Genova, Italy

[24]Sanders-Brown Center on Aging, University of Kentucky, 800 S. Limestone Street, Lexington, KY, USA

[25]Department of Biostatistics, University of Kentucky, 725 Rose Street, Lexington, KY, USA

[26]Bonn-Aachen International Center for IT, University of Bonn, Dahlmannstr. 2, Bonn, Germany

[27]Cambridge Institute for Medical Research, University of Cambridge and University of Toronto, Hills Road, Cambridge, CB2 UK

[28]McGovern Institute for Brain Research, Massachusetts Institute of Technology, 43 Vassar St, 46-4033F, Cambridge, MA, USA

[29]Department of Otology and Laryngology, Harvard Medical School, Boston, MA, USA

[30]Luxembourg Centre for Systems Biomedicine, University of Luxembourg, 7, avenue des Hauts Fourneaux, Esch-sur-Alzette, Luxembourg

[31]Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Broad Institute and Harvard Medical School, 41 Avenue Louis Pasteur, Suite 301, Boston, MA, USA

[32]Department of Computational Medicine and Bioinformatics, University of Michigan, 500 South State Street, Ann Arbor, MI, USA

[33]Department of Electrical Engineering and Computer Science, University of Michigan, 500 South State Street, Ann Arbor, MI, USA

[34]Department of Internal Medicine, University of Michigan, 500 South State Street, Ann Arbor, MI, USA

[35]Department of Biostatistics, The University of North Carolina at Chapel Hill, McGavran Greenberg Hall, CB#7422, Chapel Hill, NC, USA

[36]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Solnavägen 1, Stockholm, Sweden Paolo Inglese

[37]Department of Computer science and Statistics, Daegu University, Gyeongsan-si, Gyeongsangbuk-do, Republic of Korea

[38] Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, UK

[39]Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, UK

[40]Fraunhofer Institute for Algorithms and Scientific Computing (SCAI), Department for Bioinformatics, Schloss Birlinghoven, Sankt Augustin, Germany

[41]Computational Biology Institute, The George Washington University, 45085 University Drive, Suite 305, Ashburn, VA, USA

[42]Department of Biostatistics, Kings College London, Box P096, De Crespigny Park, SE5 8AF, London, UK

[43]Department of Neuroimmunology, Foundation Institut de Recerca, Hospital Universitari Vall d'Hebron, Ps Vall Hebron 119-129, Barcelona, Spain

[44]Department of Biology, Brigham Young University, 4102 Life Sciences Building, Brigham Young University, Provo, UT, USA

[45]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Solnavägen 1, Stockholm, Sweden

[46]Genetics and Genomics Sciences Department, Icahn School of Medicine at Moutn Sinai, 1425 Madison Ave, New York, NY, USA

[47]Genome and systems biology degree program, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd. Taipei, Taiwan

[48]Department of Radiology and Medical Imaging, The University of Virginia, Charlottesville, VA, USA

[49]Global CEO Initiative on Alzheimer's disease, 1101 K St., NW, #400, Washington, DC

[50]Radiology, Medicine, Psychiatry, and Neurology, UCSF, SFVAMC, 4150 Clement Street (114M), San Francisco, CA, USA

[51]Department of Computer Science, Hunter College, The City University of New York, 695 Park Avenue, New York, NY, USA

[52]Centre for Computational Systems Biology, Fudan University, 220 Handan Road, Shanghai, China

[53]Department of Medical Biophysics, University of Toronto, Toronto, Canada

[54]Department of Pharmacology & Toxicology, University of Toronto, Toronto, Canada

1. Abstract

Identifying accurate biomarkers of cognitive decline is essential for advancing early diagnosis and prevention therapies in Alzheimer's Disease. The Alzheimer's Disease DREAM Challenge was designed as a computational crowdsourced project to benchmark the current state-of-the-art in predicting cognitive outcomes in Alzheimer's Disease based on high-dimensional, publicly available genetic and structural imaging data. This meta-analysis failed to identify a meaningful predictor developed from either data modality, suggesting that alternate approaches should be considered for to prediction of cognitive performance.

2. Background

The Alzheimer's Disease DREAM Challenge (http://dx.doi.org/10.7303/syn2290704) was designed to provide an unbiased assessment of current capabilities for estimation of cognition and prediction of cognitive decline using genetic and imaging data from public data resources using a crowd-sourced approach. The ability to predict rate of cognitive decline – both prior to and following diagnosis – is essential to effective trial design for the development of therapies for Alzheimer's Disease (AD) prevention and treatment. Major collaborative efforts in the field are assessing the association of

genetic loci with AD diagnosis and the application of structural imaging for development of early biomarkers of diagnosis, but the utility of these approaches to estimate cognition or predict cognitive decline is not well established. This project was designed under the advisement of a panel of experts in the field to evaluate whether these questions could be meaningfully addressed with current methodologies given existing public data sources. To ensure that these questions were tested across a broad spectrum of the latest analytical approaches, the study was designed as a crowdsourced, community-based challenge in which participants were invited to address one or more of the following three problems: (1) The prediction of cognitive decline over time based on genetic data. (2) The prediction of resilience to cognitive decline in individuals with elevated amyloid burden based on genetic data. (3) The estimation of cognitive state based on structural magnetic resonance (MR) imaging data.

3. Results

3.1 Study design and data harmonization

To ensure that predictors were detecting true biological variation rather than study-specific technical variation, this project required inclusion of data from multiple study sources. While genetic and imaging data have been generated within many rich longitudinal cohorts across the field, the procurement and harmonization of these data sets was a non-trivial problem that required solutions to overcome political, ethical, and technical barriers. For example, the generation of whole genome sequencing data across multiple AD cohorts within the NIH-funded AD sequencing project has resulted in a powerful resource for genetic analysis in the field but longitudinal information on cognitive traits is not readily available in those datasets. Despite limitations on data accessibility, multiple relevant data sources were identified and used in this project including: the Alzheimer's Disease Neuroimaging Initiative (ADNI)(1), the Rush Alzheimer's Disease Center Religious Orders Study(2) and Memory and Aging Project (MAP)(3) and the European AddNeuroMed(4) study, which is part of InnoMed, a precursor to the Innovative Medicines Initiative. Data selection and processing was performed based on data availability across these three datasets. As such, cognition was defined using Mini Mental State Examination (MMSE) scores(5), genetic data was provided based on imputation across array-based genotype data, and structural MR imaging data was reprocessed in each cohort using a common processing pipeline. Genetic and imaging data was supplemented with a limited set of covariates including diagnosis, initial MMSE score, age at the initial examination, years of education, gender, and *APOE* haplotype. Participants were provided with data from ADNI to train algorithms over a four-month period and, to ensure that participation was not limited by access

to compute resources, they were offered use of the IBM z-Enterprise cloud to perform analyses. The challenge generated significant interest with 527 individuals from around the world registered to participate. A leaderboard displayed accuracy of submissions throughout the duration of the challenge: 1,157 submissions were made for problem 1,478 submissions for problem 2, and 434 submissions for problem 3. Thirty-two teams submitted final results that were scored based on prediction/estimation of blinded outcomes within ROS/MAP for genetic predictions and AddNeuroMed for imaging-based estimations (Figure 1).

3.2 Genetic prediction of cognitive decline

The first challenge question assessed the ability of current methods to predict change in cognitive examination performance based on genetic data. High prediction accuracy would signal the potential for noninvasive biomarkers of cognition to have a major clinical impact on early AD diagnosis and prevention. Previous efforts to develop predictors of change in cognitive function have not succeeded in providing robust and replicable models(6-8). Genetic variation has been demonstrated to influence AD status: rare genetic mutations at several loci are implicated in familial forms of early-onset disease(9) while common variation contributes 33% to variance in sporadic AD and 22 loci have been implicated by large-scale genetic association analyses(10, 11). However, with the exception of the *APOE*4 haplotype, there has been little success in transforming these genetic associations into meaningful clinical predictions of cognitive decline. For this purpose, participants were challenged to predict 2-year changes in MMSE scores based on genotypes imputed from SNP array data. Participants trained their algorithms with 767 ADNI samples and the algorithms' predictions were evaluated on a test set of 1,175 ROS/MAP samples with blinded outcome measures. The algorithm with the best predictive performance at the midpoint of the challenge did not contain any genetic features beyond *APOE* haplotype. Since the goal of this subchallenge was to assess genetic contribution to prediction of cognitive decline, this top-ranked algorithm was openly shared across teams as an interim baseline upon which to incorporate additional genetic predictors (http://dx.doi.org/10.7303/syn2838779). Eighteen teams submitted final predictions. The majority of methods performed significantly better than a permutation-based random model prediction (Figure 2a). A cluster of six methods performed significantly better than the others (including the interim baseline model) but were statistically indistinguishable amongst themselves (Figure 2d). Of these, the prediction with the best overall score (team GuanLab_umich from the University of Michigan) achieved a Pearson correlation of 0.382 and a Spearman correlation of 0.433 (the overall score was a rank-based combination of these two measures of performance; see online Supplement and Supplementary Methods: http://dx.doi.org/10.7303/syn3383106). However, no significant contribution

of genetics beyond *APOE* haplotype to predictive performance was observed across any of the submissions. Given the small sample size, no conclusions can be inferred from this analysis regarding the existence of genetic loci associated with cognitive decline. Rather, these observations suggest that predictors of cognitive decline developed based on genetic data will not be useful within the clinical setting.

3.3 Genetic prediction of cognitive resilience

The second question challenged participants to identify genetic predictors that could distinguish individuals who exhibit resilience to AD pathology as defined by minimal change in cognitive function despite evidence of amyloid deposition(12, 13). Identification of genetic signatures predictive of cognitive resilience would aid in the elucidation of mechanisms that may confer resilience, providing a powerful tool to help advance AD prevention strategies and treatment development. Eleven teams submitted predictions of resilience based on a training set derived from 176 ADNI subjects. Evaluations were made using data derived from 257 individuals from the ROS/MAP data. Despite using the largest such public dataset assembled to date, participants were unable to develop algorithms with predictive performances significantly better than random (see Figure 2b, online Supplement and Supplementary Methods in Synapse: http://dx.doi.org/10.7303/syn3383106). While it is likely that the study was underpowered due to small sample size and trait heterogeneity, this result suggest that information about cognitive resilience is not easily discoverable from SNP analysis.

3.4 Structural imaging-based estimation of cognition

The third question challenged participants to estimate cognitive state using structural brain image data (Figure 1, lower panel). Brain imaging has emerged as a powerful method for monitoring neurodegeneration and there is great enthusiasm in the field to make use of images for diagnosis and prediction. There have been many attempts in the past to correlate changes in brain shape with disease progression and/or diagnosis, conventionally using measures of volume for a given brain region(14, 15). More detailed shape measures of image features including cortical thickness, curvature, and depth have also been found to be relevant to a variety of neurological conditions(16). Participants were challenged to estimate MMSE scores based on structural brain images, or shape information derived from these images. Participants trained algorithms using ADNI data (N=628) and were evaluated using AddNeuromed data (N=182) for which they were blind to outcome measures. To engage as many participants as possible from both within and beyond the neuroimaging community, the data were provided both as raw MR images and as tables containing shape measures (volume, thickness, area, curvature, depth, etc.) for every labeled brain region. Thirteen teams submitted estimates for final evaluation and all teams performed better than a random model

(see online Supplement and Supplementary Methods in Synapse: http://dx.doi.org/10.7303/syn3383106). Three teams performed significantly better than the others (teams GuanLab_umich from the University of Michigan, ADDT from the Karolinska Institute and Pythia from the University of Pennsylvania) (Figure 2c) but were statistically indistinguishable from one another and tied for top average rank (Figure 2e). The algorithm that generated the best absolute mean combined rank (Team GuanLab_umich) achieved a concordance correlation coefficient of 0.569 and Pearson's correlation of 0.573 (the overall score was a rank-based combination of these two measures of performance). The most common features that contributed heavily to the MMSE estimates across the algorithms were hippocampal volume and entorhinal thickness, corroborating prior work(17-19). The top three teams also found that inclusion of shape measures of the entorhinal cortex (volume, curvature, surface area, travel and geodesic depth) improved overall estimation. Other features that contributed to predictions within the top three teams' results included volume of inferior lateral ventricle and amygdala (see online Supplement and Supplementary Methods in Synapse: http://dx.doi.org/10.7303/syn3383106). These results validate an established relationship between structural imaging data and cognition. However, the correlative performance of these estimators was low suggesting that their application in the clinical setting may not be sufficient to inform patient care.

4. Discussion

The AD DREAM Challenge provided a formalized assessment of the ability to develop meaningful predictions of cognitive performance from public genetic or imaging data using contemporary state-of-the-art predictive algorithms. Predictive performance across all three of the subchallenges was modest and most methods performed roughly equivalently. Given this uniform performance, we do not expect that the presented results are a failure of current modeling methodologies. A more likely explanation is that the data used to address these questions were inadequate to support these tasks. We also note that the majority of research teams that participated in this challenge did not have expertise in the field of AD. Although the few teams that did posses this knowledge did not do better than the others, there remains the possibility that performance would have been improved by the inclusion of more domain experts.

4.1 Use of genetic information for cognitive prediction

The modest performance observed in the subchallenges focused on genetic analysis demonstrated that contemporary algorithms were not able to leverage genetic signal to make useful predictions for cognition.   These results support the prevailing expectation that genetic variants of moderate to high

frequency will not support viable biomarker development in AD (9-11). Although heritability estimates and linkage studies have demonstrated that there is a considerable estimated genetic contribution to AD onset and progression (11, 20, 21), evidence both within the AD field and across other complex disease (22) traits has indicated that this overall genetic contribution is the aggregated compilation of a large number of loci with small – independent or epistatic – effects. Historically, this type of signal is difficult to capture in predictive models and unlikely to be useful in a diagnostic setting (23). Furthermore, cognition is highly influenced by a host of non-genetic factors relating to lifestyle choices and accumulated exposures that were not represented across all of these datasets and, in fact, are not fully captured in most cohorts (24-27). Non-genetic contributions to cognitive performance may themselves provide an important base for successful predictions. Within the context of genetic analysis, the absence of these factors from models confounds the ability to detect real genetic signal and impacts the ability to accurately model state-specific genetic contributions. As such, future inquiry into the use of genetic testing for prediction of cognitive performance and AD risk assessment may be better served by focusing on the contribution of rare genetic variation. Recently discovered disease-associated rare variants have larger effect sizes than common variants and confer 2 to 5 fold greater risk or protection in carriers relative to the general population (28-30). Ongoing large-scale sequencing analyses will identify additional associated rare risk variants. In sufficient numbers, the aggregate prevalence would support the development of a genetic diagnostic containing a library of rare variants.

4.2 Use of structural imaging data for cognitive estimation

While the inexpensive and noninvasive nature of genetic testing makes this approach amenable to population-level disease screening, the resource-intensive nature of image-based testing is better positioned for careful evaluation of high-risk individuals. As such, these approaches are needed to provide a higher confidence estimate of cognitive performance. Although a variety of methods developed within the context of this challenge were able to successfully estimate cognition, none of these methods were sufficiently accurate to merit clinical consideration. These observations support previous work in the field (17, 19) and highlight the imperfect relationship between brain structure and function. Newer imaging modalities that focus on brain function and/or pathology – such as FDG-PET (31) or tau imaging (32)– may prove more successful for assessing cognitive dysfunction.

4.3 Effective performance of meta-analysis across diverse cohorts

A major consideration for any meta-analysis is the issue of appropriate harmonization of data across disparate sources.   Despite leveraging several of the most deeply phenotyped cohorts in the field, this challenge limited analysis to those traits that were in common across cohorts. Although this

approach to data harmonization is standard practice for meta-analyses (10), it greatly reduced the depth of the information available for modeling and influenced the selection of cognitive measures for use as prediction outcomes. Because each cohort had performed a battery of study-specific tests, this greatly limited the ability for finer grained assessment across cognitive processes. A more sensible approach for future analyses may be to focus effort on more sophisticated methods to calibrate disparate cognitive phenotypes across cohorts (33).   Another undesirable consequence of the focus on traits measured in common was the inability to incorporate into model development the full spectrum of non-genetic and non-imaging factors that are known to influence cognitive performance (24-27). This suggests the need for development of alternate approaches for integrating heterogeneous data and/or assessing replication across cohorts. Alternatively, smaller scale analyses that prioritize phenotypic depth over sample size may afford a more refined view of disease.

In summary, this challenge demonstrated that predictions of cognitive performance developed from genetic or structural imaging data were modest across a diverse set of contemporary modeling methods. Future efforts to identify clinically relevant predictors of cognition will benefit from a focus on alternate data sources as well as methods that work to incorporate greater phenotypic complexity.

FIGURE LEGENDS:

Fig.1. Challenge overview.
The top schematic summarizes the three challenge questions on the left column, the training data in the middle, and the test data on the right, including numbers of subjects. The symbols represent sources of data (demographic, ROS/MAP genetic, and ADNI or ANM brain images and shape information). The bottom panel provides example brain image labels and shape information provided to the participants for question 3. Anatomical labels for left cortical regions are shown on the left and just a couple of the cortical surface shape measures are shown on the right (travel depth on top and mean curvature below), for both uninflated and inflated surfaces (top and bottom rows, respectively).

Figure 2. Performance evaluation results. Panels a, b, and c report the p-values (in negative log10 scale) for intersection union tests investigating which teams performed better than random for questions 1, 2, and 3, respectively. Explicitly, for question 1 (panel a) we tested the null hypothesis that at least one of the four correlation coefficients (namely, Pearson/clinical, Pearson/clinical + genetics, Spearman/clinical, Spearman/clinical + genetics) is equal to zero, against the alternative that all four correlation coefficients are larger than zero. Adopting a 0.05 significance level, 26 out of the 32 submissions were statistically better than random, after Bonferroni multiple testing correction for 32 tests (submissions crossing the black vertical line). For question 2 (panel b), we tested the null hypothesis that balanced accuracy = 0.5 or AUC = 0.5, against the alternative that balanced accuracy > 0.5 and AUC > 0.5. In this case, no model performed significantly better than random and, therefore, no best performer was declared. For question 3 (panel c), we tested the null hypothesis that Pearson's correlation (COR) or Lin's concordance correlation coefficient (CCC) are equal to zero, against the alternative that both COR and CCC are larger than zero. Adopting a 0.05 significance level, all 23 submissions were statistically better than random, after Bonferroni correction. For all three questions, the p-values were computed from an empirical null distribution based on 10,000 permutations. Panels d and e report the bootstrapped assessment of ranks for questions 1 and 3, respectively. Samples were resampled with replacement from the original data (true outcome and team's predictions), and the ranks of the different teams were re-assessed in each of 100,000 re-samplings. Submissions were sorted according to the median of their bootstrapped average ranking distributions. The black horizontal line represents the posterior odds cutoff from the Bayesian

analysis. Teams above the black line are statistically tied to the top ranked model, according to a posterior odds threshold of 3.

REFERENCES:

[1]   Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack C, Jagust W, et al. The Alzheimer's disease neuroimaging initiative. Neuroimaging clinics of North America. 2005 Nov;15(4):869-77, xi-xii. PubMed PMID: 16443497. Pubmed Central PMCID: 2376747.
[2]   Bennett DA, Schneider JA, Arvanitakis Z, Wilson RS. Overview and findings from the religious orders study. Current Alzheimer research. 2012 Jul;9(6):628-45. PubMed PMID: 22471860. Pubmed Central PMCID: 3409291.
[3]   Bennett DA, Schneider JA, Buchman AS, Barnes LL, Boyle PA, Wilson RS. Overview and findings from the rush Memory and Aging Project. Current Alzheimer research. 2012 Jul;9(6):646-63. PubMed PMID: 22471867. Pubmed Central PMCID: 3439198.
[4]   Lovestone S, Francis P, Kloszewska I, Mecocci P, Simmons A, Soininen H, et al. AddNeuroMed--the European collaboration for the discovery of novel biomarkers for Alzheimer's disease. Annals of the New York Academy of Sciences. 2009 Oct;1180:36-46. PubMed PMID: 19906259.
[5]   Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. Journal of psychiatric research. 1975 Nov;12(3):189-98. PubMed PMID: 1202204.
[6]   Ercoli LM, Siddarth P, Dunkin JJ, Bramen J, Small GW. MMSE items predict cognitive decline in persons with genetic risk for Alzheimer's disease. Journal of geriatric psychiatry and neurology. 2003 Jun;16(2):67-73. PubMed PMID: 12801154.
[7]   Hsiung GY, Alipour S, Jacova C, Grand J, Gauthier S, Black SE, et al. Transition from cognitively impaired not demented to Alzheimer's disease: an analysis of changes in functional abilities in a dementia clinic cohort. Dementia and geriatric cognitive disorders. 2008;25(6):483-90. PubMed PMID: 18417973.
[8]   Vemuri P, Wiste HJ, Weigand SD, Shaw LM, Trojanowski JQ, Weiner MW, et al. MRI and CSF biomarkers in normal, MCI, and AD subjects: predicting future clinical change. Neurology. 2009 Jul 28;73(4):294-301. PubMed PMID: 19636049. Pubmed Central PMCID: 2715214.
[9]   Ridge PG, Ebbert MT, Kauwe JS. Genetics of Alzheimer's disease. BioMed research international. 2013;2013:254954. PubMed PMID: 23984328. Pubmed Central PMCID: 3741956.
[10]   Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nature genetics. 2013 Dec;45(12):1452-8. PubMed PMID: 24162737. Pubmed Central PMCID: 3896259.
[11]   Ridge PG, Mukherjee S, Crane PK, Kauwe JS, Alzheimer's Disease Genetics C. Alzheimer's disease: analyzing the missing heritability. PloS one. 2013;8(11):e79771. PubMed PMID: 24244562. Pubmed Central PMCID: 3820606.

[12]   Bennett DA, Schneider JA, Arvanitakis Z, Kelly JF, Aggarwal NT, Shah RC, et al. Neuropathology of older persons without cognitive impairment from two community-based studies. Neurology. 2006 Jun 27;66(12):1837-44. PubMed PMID: 16801647.

[13]   Price JL, Morris JC. Tangles and plaques in nondemented aging and "preclinical" Alzheimer's disease. Annals of neurology. 1999 Mar;45(3):358-68. PubMed PMID: 10072051.

[14]   Davatzikos C, Xu F, An Y, Fan Y, Resnick SM. Longitudinal progression of Alzheimer's-like patterns of atrophy in normal older adults: the SPARE-AD index. Brain : a journal of neurology. 2009 Aug;132(Pt 8):2026-35. PubMed PMID: 19416949. Pubmed Central PMCID: 2714059.

[15]   Misra C, Fan Y, Davatzikos C. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI. NeuroImage. 2009 Feb 15;44(4):1415-22. PubMed PMID: 19027862. Pubmed Central PMCID: 2648825.

[16]   Im K, Lee JM, Seo SW, Hyung Kim S, Kim SI, Na DL. Sulcal morphology changes and their relationship with cortical thickness and gyral white matter volume in mild cognitive impairment and Alzheimer's disease. NeuroImage. 2008 Oct 15;43(1):103-13. PubMed PMID: 18691657.

[17]   Haight TJ, Jagust WJ, Alzheimer's Disease Neuroimaging I. Relative contributions of biomarkers in Alzheimer's disease. Annals of epidemiology. 2012 Dec;22(12):868-75. PubMed PMID: 23102709. Pubmed Central PMCID: PMC3510749.

[18]   Nho K, Risacher SL, Crane PK, DeCarli C, Glymour MM, Habeck C, et al. Voxel and surface-based topography of memory and executive deficits in mild cognitive impairment and Alzheimer's disease. Brain imaging and behavior. 2012 Dec;6(4):551-67. PubMed PMID: 23070747. Pubmed Central PMCID: 3532574.

[19]   Thung KH, Wee CY, Yap PT, Shen D, Alzheimer's Disease Neuroimaging I. Neurodegenerative disease diagnosis using incomplete multi-modality data via matrix shrinkage and completion. NeuroImage. 2014 May 1;91:386-400. PubMed PMID: 24480301. Pubmed Central PMCID: 4096013.

[20]   Escott-Price V, Sims R, Bannister C, Harold D, Vronskaya M, Majounie E, et al. Common polygenic variation enhances risk prediction for Alzheimer's disease. Brain : a journal of neurology. 2015 Dec;138(Pt 12):3673-84. PubMed PMID: 26490334.

[21]   Lee SH, Harold D, Nyholt DR, Consortium AN, International Endogene C, Genetic, et al. Estimation and partitioning of polygenic variation captured by common SNPs for Alzheimer's disease, multiple sclerosis and endometriosis. Hum Mol Genet. 2013 Feb 15;22(4):832-41. PubMed PMID: 23193196. Pubmed Central PMCID: PMC3554206.

[22]   Chatterjee N, Wheeler B, Sampson J, Hartge P, Chanock SJ, Park JH. Projecting the performance of risk prediction based on polygenic analyses of genome-wide association studies. Nature genetics. 2013 Apr;45(4):400-5, 5e1-3. PubMed PMID: 23455638. Pubmed Central PMCID: PMC3729116.

[23]   Manolio TA. Bringing genome-wide association findings into clinical use. Nat Rev Genet. 2013 Aug;14(8):549-58. PubMed PMID: 23835440.

[24]   Scarmeas N, Stern Y, Mayeux R, Luchsinger JA. Mediterranean diet, Alzheimer disease, and vascular mediation. Arch Neurol. 2006 Dec;63(12):1709-17. PubMed PMID: 17030648. Pubmed Central PMCID: PMC3024906.

[25]   Podewils LJ, Guallar E, Kuller LH, Fried LP, Lopez OL, Carlson M, et al. Physical activity, APOE genotype, and dementia risk: findings from the Cardiovascular Health Cognition Study. Am J Epidemiol. 2005 Apr 1;161(7):639-51. PubMed PMID: 15781953.

[26]   Lindsay J, Laurin D, Verreault R, Hebert R, Helliwell B, Hill GB, et al. Risk factors for Alzheimer's disease: a prospective analysis from the Canadian Study of Health and Aging. Am J Epidemiol. 2002 Sep 1;156(5):445-53. PubMed PMID: 12196314.

[27]   Wang HX, Karp A, Winblad B, Fratiglioni L. Late-life engagement in social and leisure activities is associated with a decreased risk of dementia: a longitudinal study from the Kungsholmen project. Am J Epidemiol. 2002 Jun 15;155(12):1081-7. PubMed PMID: 12048221.

[28] Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeva E, Majounie E, et al. TREM2 variants in Alzheimer's disease. N Engl J Med. 2013 Jan 10;368(2):117-27. PubMed PMID: 23150934. Pubmed Central PMCID: PMC3631573.

[29] Jonsson T, Stefansson H, Steinberg S, Jonsdottir I, Jonsson PV, Snaedal J, et al. Variant of TREM2 associated with the risk of Alzheimer's disease. N Engl J Med. 2013 Jan 10;368(2):107-16. PubMed PMID: 23150908. Pubmed Central PMCID: PMC3677583.

[30] Jonsson T, Atwal JK, Steinberg S, Snaedal J, Jonsson PV, Bjornsson S, et al. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. Nature. 2012 Aug 2;488(7409):96-9. PubMed PMID: 22801501.

[31] Gray KR, Wolz R, Heckemann RA, Aljabar P, Hammers A, Rueckert D, et al. Multi-region analysis of longitudinal FDG-PET for the classification of Alzheimer's disease. NeuroImage. 2012 Mar;60(1):221-9. PubMed PMID: 22236449. Pubmed Central PMCID: PMC3303084.

[32] James OG, Doraiswamy PM, Borges-Neto S. PET Imaging of Tau Pathology in Alzheimer's Disease and Tauopathies. Front Neurol. 2015;6:38. PubMed PMID: 25806018. Pubmed Central PMCID: PMC4353301.

[33] Gross AL, Sherva R, Mukherjee S, Newhouse S, Kauwe JS, Munsie LM, et al. Calibrating longitudinal cognition in Alzheimer's disease across diverse test batteries and datasets. Neuroepidemiology. 2014;43(3-4):194-205. PubMed PMID: 25402421. Pubmed Central PMCID: PMC4297570.