A blood test for Alzheimer's disease: progress, challenges and recommendations

Steven J. Kiddle[a,b,*], Nicola Voyle[a], Richard JB Dobson[a,c,d,*]

[a] Department of Biostatistics & Health Informatics, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK.

[b] MRC Biostatistics Unit, Cambridge Institute of Public Health, University of Cambridge, Cambridge, UK.

[c] NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at South London and Maudsley NHS Foundation, London, UK.

[d] Farr Institute of Health Informatics Research, UCL Institute of Health Informatics, University College London, London, UK.

* Correspondence:

Dr Steven Kiddle, +44 (0)1223 760717, steven.kiddle@mrc-bsu.cam.ac.uk

OR

Professor Richard Dobson, +44 (0)20 7848 0924, richard.j.dobson@kcl.ac.uk

Running title: A blood test for Alzheimer's disease

Keywords: Alzheimer disease, **blood tests**, **cohort studies**, **research design**, **data reporting**, **genetics**, **blood proteins**, **gene expression**, **metabolomics**

Abstract

Ever since the discovery of *APOE ε4* around 25 years ago researchers have been excited about the potential of a blood test for Alzheimer's disease (AD). Since then researchers have looked for genetic, protein, metabolite and/or gene expression markers of AD and related phenotypes. However, no blood test for AD is yet being used in the clinical setting. We first review the trends and challenges in AD blood biomarker research, before giving our personal recommendations to help researchers overcome these challenges. While some degree of consistency and replication has been seen across independent studies, several high-profile studies have seemingly failed to replicate. Partly due to academic incentives there is a reluctance in the field to report predictive ability, to publish negative findings and to independently replicate the work of others. If this can be addressed then we will know sooner whether a blood test for AD or related phenotypes with clinical utility can be developed.

Progress

The identification of genetic markers such as *APOE ε4* arguably represented the first step change in progress towards a blood test for late onset Alzheimer's disease (AD), as genetic markers can be measured from blood samples [1]. Since then 19 other significant markers of AD have been identified by a Genome-Wide Association Study (GWAS) [2]. These markers have been combined into a polygenic risk score with ~80,000 more weakly associated genetic markers achieving an Area Under the Curve (AUC) of 78% for prediction of AD. This compares with 72% achievable with just age, sex and *APOE ε4* (the `co-variate only' model) [3]. In a smaller recent study (N ~

2

1,600) some of the same authors have shown that the same risk score has an AUC of 84% for predicting pathologically confirmed cases [4], which if confirmed in larger studies may have enough clinical utility to justify the use of genome-wide genotyping in the clinic to aid diagnosis. Even an AUC of 78%, if validated further, may have some utility for recruitment of higher risk individuals to prevention trials [5]. Other promising approaches at an earlier stage of development have involved increasing the GWAS sample size using an AD-by-proxy phenotype [6], and developing polygenic hazard scores to predict age of dementia onset [7].

AD polygenic risk scores have also been shown to be associated with brain atrophy [8,9] and cerebrospinal fluid (CSF) amyloid beta [7,8], but to the best of our knowledge they have not yet been shown to be predictive of these phenotypes. In fact, two independent studies (N = 657 and N = 242) have shown that this genetic risk score does not appear to be predictive of amyloid or tau pathology measured from CSF [10,11]. A preliminary report that the polygenic hazard score is better able to predict elevated brain amyloid does not appear to test whether it improves upon a model using age, gender and APOE alone, so its clinical utility is uncertain [12]. If the negative findings are correct then this would be consistent with the idea that biomarkers must be optimised to specific use contexts, e.g. useful markers of AD diagnosis may differ from markers of pathology or progression. For this reason the literature has broadened out from the case-control design to encompass endophenotype designs which look for markers of brain atrophy [13], CSF pathology [14] or cognitive decline [15]. However, it should be noted these findings do not yet seem to have translated into accurate prediction models [16].

Following on from the early genetic work, proteomic researchers joined the search for blood biomarkers of AD. In a clear parallel to earlier genetics research two main approaches were taken: 'candidate' studies and 'discovery' studies. The AlzBiomarker project has performed a meta-analysis of association studies of 'candidate' blood (and CSF) markers with AD, showing that of all candidates in blood, total-tau appears to be the most associated with AD [17]. Predictive results from the use of blood total-tau are only provided for individual small-scale studies, so the reported sensitivity and specificity for predicting AD diagnosis is likely to be overly optimistic, e.g. 97% specificity and 91% sensitivity in Chiu et al., [18].

We have reviewed the 21 'discovery' blood protein studies published between 2002 - 2014, using a wide-range of proteomics techniques and AD-related outcomes. A low consistency of biomarkers identified between studies was observed, but four candidate biomarkers were observed in studies utilizing five independent research cohorts: α-1-antitrypsin, α-2-macroglobulin, apolipoprotein E, and complement C3. When examined in a new dataset these proteins, when combined with age, sex and presence of *APOE $\varepsilon4$*, had an AUC of 82% for predicting AD diagnosis, versus 79% for the co-variate only model [19]. While superficially better than the genetic risk scores AUC (78% versus 82%), it should be noted that the genetic result is more trustworthy as it comes from a much larger study and shows a greater difference in predictive ability between co-variate only and biomarker models.

In another parallel to genetic research, proteomic studies have also looked-for biomarkers of endophenotypes of AD, which have been reviewed in Baird et al., [20] who see promise in this area, but acknowledge limited success in identification of a reproducible signature. More recently,

Nakamura et al., [21] have developed a blood test based on fragments of the protein amyloid beta that achieved an AUC of 94% for the prediction of elevated brain amyloid in a moderately sized study (N =232). If shown to be robust, reliable, practical and affordable this could be a major breakthrough.

Blood metabolite studies of AD are more novel, with high profile papers by Mapstone et al., [22] and Proitsi et al., [23]. Mapstone et al., [22] identified ten lipids which could predict conversion from Mild Cognitive Impairment to AD over 2 - 3 years with an AUC of 92%. Proitsi et al., [23] identified 24 metabolites which had an AUC of 71% for predicting AD diagnosis, in a considerably larger study (N = 277 versus N = 85).

Gene expression has also been explored as a potential source of blood biomarkers for AD. Little consistency has been seen in the genes selected by these various studies, leading Han et al., [24] to suggest that greater concordance might be seen at the pathway level. We demonstrated in Voyle et al., [25] a failure to replicate classifiers between independent sample sets, and that simple pathway level summaries of gene expression are no more predictive of AD. Endophenotype approaches have been explored in gene expression studies as well, for example Lunnon et al., [26] show that gene expression is predictive of brain atrophy. However, this work has not yet been replicated in independent cohorts.

More recently, we have been attempting to combine different modalities of biomarker to improve predictive ability. In Voyle et al., [27] we found that five metabolites could be used to predict amyloid positive individuals with 72% accuracy, rising to 79% when combined with levels of the

protein fibrinogen gamma. This study was limited in size and requires replication in independent samples. In a similar vein, we showed a marginal improvement in the prediction of CSF amyloid beta levels using genetic risk from the large AD GWAS [2,3] and plasma tau levels (AUC 67% versus 66% for co-variate only model) [10]. Such a small improvement may be artefactual, and even if true is not likely to be useful by itself. Studies seeking to find multi-modal AD blood biomarkers and/or biomarkers of endophenotypes are likely to become more common but have been held back by the sample sizes available, which is a focus for improvement going forward.

Challenges and recommendations

Further progress towards a clinically useful blood test will be slow unless we acknowledge and learn from the limitations of our current approaches, therefore what follows is our personal view on key challenges and important recommendations for future AD blood biomarker research. We draw attention to limitations of existing studies not to dismiss them, but to point out room for improvement in the field and in our own research. What follows will seem obvious to some, but needs to be highlighted as a counter-point to the over-optimism of the field.

The quality of experimental design in this field is variable, although this has been improving over time. One of the most obvious aspects of this is in sample size of non-genetic AD biomarker discovery studies, for example a study seeking to find blood protein AD biomarkers in 2002 used only 18 research participants [28]. Ten years later Doecke et al., [29] achieved N ~ 1000. This seemed a positive trend, but unfortunately, we are not aware of any larger studies published in the five years that followed. In fact, many smaller-scale studies (N ~ 100) are still published, e.g.

[22,27,30,31]. We should learn from the field of genetics, where small-scale candidate gene studies were plagued with replication issues [32] that were only solved by larger sample size and unbiased approaches. We recommend that this is tackled, in part, using samples from larger cohorts, such as UK Biobank [33] and the Precision Medicine Initiative [34], as well as cohort consortia such as the European Medical Information Framework - AD (http://www.emif.eu/) and Dementia Platform UK (http://www.dementiasplatform.uk).

Additional design considerations involve the appropriateness of the population used, this should be guided by the anticipated context of use of the potential blood tests. Most studies have sought to find a blood test that could be helpful in the diagnosis of Alzheimer's disease, however none have yet been performed in the primary healthcare population in which it would have greatest utility [35]. This has been in large part due to the challenges of recruiting research participants, and the priority given to large scale recruitment rather than to representativeness of populations relative to anticipated context-of-use. Similarly, studies have sought to find AD markers that could be used to identify asymptomatic patients with early signs of AD, but very few have been performed in that population [36,37,38,39]. Another problem of unrepresentative sample populations is that they may not reflect the prevalence of AD related phenotypes (e.g. amyloid positivity) in populations appropriate to the anticipated context-of-use, which could inflate the positive and negative predictive values.

Partly due to the history of this field, which was initially led by clinicians and lab biologists, the level of statistical rigour is understandably variable. One example is the focus of many papers on p-values instead of predictive ability (sensitivity/specificity/positive predictive value etc).

Significant p-values, even if replicable, do not necessarily mean that a biomarker is useful for predicting AD. To do so requires a suitably large effect size and a good understanding of confounding factors (e.g. age, gender, *APOE ε4*, medication use). For readers who may struggle to interpret and understand predictive measures we heartily recommend Tze-Wey Loong's excellent visual explanation [40].

A critically important consideration is cross-validation, i.e. the assessment of predictive models in additional data not used in its construction. The data used in model construction is referred to as the training or in-sample dataset, whereas the independent data used for assessment is called the test or extra-sample data. It is important that predictive performance is reported from the test data, as results from training data can be artefactually better due to overfitting to noise [41]. This is equivalent in importance to blinding in clinical trials in the sense that it helps to protect results from the preconceptions of the researcher. This can be a major concern in studies which report predictive accuracy in training sets only, i.e. where no cross-validation has been performed. In some studies, k-fold cross-validation has been performed, in which the training data is repeatedly split into different training and test subsets and average performance in the test sets given. This is better than no cross-validation, but can still give overly generous predictive performance due to the train and test datasets sharing the same systematic noise, especially in small datasets and when $k < 10$ [42].

Despite seemingly performing cross-validation, researchers can often subconsciously and artefactually inflate predictive performance. The two most common examples of this are (1) when both train and test data are used for variable selection, for example when variables are selected for

model inclusion by ranking the p-values from repeated univariate tests using all available data (and therefore not correctly holding out test data), and (2) by performing correct cross-validation, finding a poor result which does not get reported, and then trying a new model (new variables, or new modelling approach or formula), only reporting models which have good predictive performance. The former mistake is one that we have made ourselves in Kiddle et al., [30], the latter mistake is probably the most common which is evidenced in part by the relative absence of negative results published in this area. The absence of negative results is a form of reporting bias, leading to an overly optimistic impression of this field in the literature. While no evidence of reporting bias for 'candidate' blood protein markers of AD is seen in the AlzBiomarker meta-analysis, in the context of discovery and replication of blood biomarker panels we are only aware of the following relevant papers showing negative results, some emerging from our own group [10,25], some consistent with our negative findings [11], and others showing limited [43,44] or no replication [45] of previous high-profile positive results [22,46] when examined by independent researchers. It seems implausible that negative results are truly rare in AD blood biomarker research.

In terms of root causes, we are generally dis-incentivised to share negative or disappointing results, even if they are correct. However, the examples provided above demonstrate that it is simply not true that journals refuse to publish negative results, although we have certainly argued against peer reviewers who fail to see the value in doing so. In a worst case scenario negative results can be published on pre-print servers such as bioRxiv (http://biorxiv.org/). It may be worthwhile for the community to come together to generate a pre-registration platform, and to persuade journals to only publish pre-registered AD biomarker studies, so that groups failing to publish negative

findings can be pressured into doing so. While not explicitly focused on blood markers, a very positive move in this direction has been challenges in which predictions are submitted before test data is released, notably the AD DREAM challenge [16], which generated a negative result, and the TADPOLE challenge which only recently stopped accepting submissions (https://tadpole.grand-challenge.org/).

Another cause could be confirmation bias, the psychological phenomena where individuals seek out evidence that matches their pre-conceived ideas [47]. This can mean that if a researcher believes a blood marker of AD will be found they may disregard evidence to the contrary. This can lead to Hypothesising After Results Are Known, researchers testing many different hypothesis (or models) reporting only the positive results and ignoring the multiple testing problem relating to all the unreported negative results [48]. This has also been called the file drawer effect, meaning that many published results are false positives [49]. John Ioannidis has discussed this in his controversially titled paper – "why most published research findings are false" [50]. While this focuses on p-values, it is likely to apply to studies reporting predictive performance as well. Interestingly, the field of blood biomarkers for AD fits his criteria for risk of high false positive publication rate: (1) small studies, (2) small effect sizes, (3) greater number and lesser pre-selection of tested relationships, (4) flexibility in designs and analyses, (5) financial interests and prejudice, and (6) many teams involved in a scientific field in chase of statistical significance.

The testing of promising biomarkers by independent researchers is essential to progress them towards the clinic [35], however this is rare. This has therefore been a focus of our research [19,51]. Where this have been done for two high-profile studies [22,46] they show a complete failure to

replicate [45] or a significantly less promising performance [43,44]. Less promising replications fit with the findings of John Ioannidis [52] who has studied the reasons that discovery results are typically inflated. In terms of clinical utility a high predictive performance is required, meaning that partial replications in well-designed studies almost always rule out the clinical utility of the marker.

Given all the above it is healthy to be skeptical about the potential of biomarkers in the current literature. It is a safe bet to assume that existing AD biomarker candidates at the very least require further validation and at worst are non-replicable, or are not of sufficiently high performance for clinical use. We hope this changes soon, but we believe that our recommendations may help the field to achieve this sooner.

While it is tempting to think that these problems may plague all blood biomarker research, this can be disproved by the example of blood protein markers of aging. Using proteomics techniques that have also been used in AD research we have discovered plasma proteins correlating highly with chronological age, that replicate strongly in an independent cohort [53], in another cohort in serum [50] and in at least two studies by independent researchers [54,55]. This shows that failures to replicate AD biomarkers cannot solely be blamed on the measurement technologies used, but it is certainly true that if a novel technology was able to detect a stronger AD related signal then it is more likely to be replicated.

Given the large reproducibility crisis in science [50], and in this field specifically, how do we improve the way we do research to increase the chance that a reproducible blood test can be found?

John Ioannidis has provided general recommendations for researchers to improve the chance of true positive findings that we think are relevant: "large-scale collaborative research, replication culture, registration, sharing, reproducibility practices, better statistical methods, standardization of definitions and analyses, more appropriate (usually more stringent) statistical thresholds, and improvement in study design standards, peer review, reporting and dissemination of research, and training of the scientific workforce" [56]. Other recommendations from John Ioannidis regard institutional changes affecting the incentives for scientists which would be even harder to achieve.

As highlighted by John Ionnadis sharing and quality of reporting is important, and for that reason we strongly recommend that the field adopt the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) reporting guidelines [57]. Specifically for the reporting of pre-analytical variables, we recommend the advice provided by Sid O'Bryant et al., [58].

In terms of sharing we recommend that raw data should be shared as widely as possible, allowing independent statisticians to verify reproducibility of findings. This has been most successfully achieved through the Alzheimer's Disease Neuroimaging Initiative (https://ida.loni.usc.edu/), but is also being done to some extent by other cohorts including the Australian Imaging and Behaviour Longitudinal study of aging (https://ida.loni.usc.edu/) and AddNeuroMed (https://www.synapse.org/#!Synapse:syn4907804). Clinicians leading large data collection need to be won over by the benefits of data sharing, including by inclusion of appropriate authorships (e.g. of clinical consortia), to counter the recent high-profile accusation that data sharing leads to "research parasites" [59]. While data sharing is done to some extent, the field would be greatly

improved by researchers also sharing analysis scripts wherever possible. This would allow errors to be spotted, would make reproducibility straightforward, and would greatly assist junior researchers to develop the coding skills that are increasingly important for modern biomedical research. Code sharing can be facilitated by websites such as GitHub (https://github.com/), which allow both private and public sharing of code, as well as version control. Care must of course be taken not to release data that is sensitive, including within the script itself, but we have shown that this can be achieved [60].

Our final recommendation is that the ultimate aim is prediction models useful in a given clinical context, and that we shouldn't be limiting ourselves to looking for markers in blood. Other variables, derived from routine clinical data such as electronic health records, wearables or cognitive tests may have more promise for this purpose. This is not to say analysis of these datasets is not without its own challenges. Electronic health records represent sparse, incomplete and often subjective representations of the disease state with important data often buried within free text narrative. The opportunities here are vast though, and although the datasets represent secondary use data, the data themselves are much larger and more representative of clinical settings than we are typically used to in blood biomarker studies, albeit less controlled in terms of co-variates and missing data. We have established research programmes in this area and through information and extraction toolkits such as Clinical Records Interactive Search [61], CogStack [62] and the KConnect programme (KConnect.eu), rolled out to multiple hospitals. We have used natural language processing to explore questions that include characterising trajectories of cognitive decline with a specific focus on identifying and validating associations, with medications for example [63]. The ubiquitous use of smartphones and wearables devices provides the opportunity

for a more objective, continuous and pervasive phenotype, throughout the disease continuum from at risk, early diagnosis through to post diagnosis engagement, compliance and self-management. Such data provides the opportunity to augment our blood biomarker studies and clinical trials. We have established programmes such as the RADAR-CNS (RADAR-CNS.org), a major goal of which is to develop a generalised real-time streaming platform that will enable active (eg through questionnaires) and passive (eg. accelerometry and heart rate using sensors on wearables and devices) remote monitoring, tracking phenotypes such as function and cognition.

Conclusions

The field of AD blood biomarkers has expanded to include genetic, protein, metabolite and gene expression markers, as well as combinations of the above. While some consistency has been seen across independent studies, several high-profile studies have seemingly failed to replicate. At the same time, there appears to be a strong reporting bias for studies seeking to find a biomarker panel with few negative results published, making the true state of play impossible to assess. Will a clinically useful blood test for AD be developed? It is simply too early to say, but we will have a better chance if we can improve the design, analysis and reporting of studies. Many alternative markers exist within health records, from wearables or innovative cognitive tests, and these should also be explored.

Acknowledgements

References

[1]   Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, Roses AD, Haines JL, Pericak-Vance MA (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* **261** (5123), 921-923.

[2]   Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, Jun G, DeStefano AL, Bis JC, Beecham GW, Grenier-Boley B, Russo G, Thornton-Wells TA, Jones N, Smith AV, Chouraki V, Thomas C, Ikram MA, Zelenika D, Vardarajan BN, Kamatani Y, Lin CF, Gerrish A, Schmidt H, Kunkle B, Dunstan ML, Ruiz A, Bihoreau MT, Choi SH, Reitz C, Pasquier F, Hollingworth P, Ramirez A, Hanon O, Fitzpatrick AL, Buxbaum JD, Campion D, Crane PK, Baldwin C, Becker T, Gudnason V, Cruchaga C, Craig D, Amin N, Berr C, Lopez OL, De Jager PL, Deramecourt V, Johnston JA, Evans D, Lovestone S, Letenneur L, Moron FJ, Rubinsztein DC, Eiriksdottir G, Sleegers K, Goate AM, Fiévet N, Huentelman MJ, Gill M, Brown K, Kamboh MI, Keller L, Barberger-Gateau P, McGuinness B, Larson EB, Green R, Myers AJ, Dufouil C, Todd S, Wallon D, Love S, Rogaeva E, Gallacher J, George-Hyslop PS, Clarimon J, Lleo A, Bayer A, Tsuang DW, Yu L, Tsolaki M, Bossù P, Spalletta G, Proitsi P, Collinge J, Sorbi S, Sanchez-Garcia F, Fox NC, Hardy J, Naranjo MCD, Bosco P, Clarke R, Brayne C, Galimberti D, Mancuso M, Matthews F, Moebus S, Mecocci P, Del Zompo M, Maier W, Hampel H, Pilotto A, Bullido M, Panza F, Caffarra P, Nacmias B, Gilbert JR, Mayhaus M, Lannfelt L, Hakonarson H, Pichler S, Carrasquillo MM, Ingelsson M, Beekly D, Alvarez V, Zou F, Valladares O, Younkin SG, Coto E, Hamilton-Nelson KL, Gu W, Razquin C, Pastor P, Mateo I, Owen MJ, Faber KM, Jonsson PV, Combarros O, O'Donovan MC, Cantwell LB, Soininen H, Blacker D, Mead S, Mosley TH, Bennett DA, Harris TB, Fratiglioni L, Holmes C, de Bruijn RFAG, Passmore P, Montine TJ, Bettens K, Rotter JI, Brice A, Morgan K, Foroud TM, Kukull WA, Hannequin D, Powell J F, Nalls MA, Ritchie K, Lunetta KL, Kauwe JSK, Boerwinkle E, Riemenschneider M, Boada M, Hiltunen M, Martin ER, Schmidt R, Rujescu D, Wang LS, Dartigues JF, Mayeux R, Tzourio C, Hofman A, Nöthen MM, Graff C, Psaty BM, Jones L, Haines JL, Holmans PA, Lathrop M, Pericak-Vance MA, Launer LJ, Farrer LA, van Duijn CM,

Van Broeckhoven C, Moskvina V, Seshadri S, Williams J, Schellenberg GD, Amouyel P (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **45**, 1452–1458.

[3] Escott-Price V, Sims R, Bannister C, Harold D, Vronskaya M, Majounie E, Badarinarayan N, Morgan K, Passmore P, Holmes C, Powell J, Brayne C, Gill M, Mead S, Goate A, Cruchaga C, Lambert J C, Van Duijn C, Maier W, Ramirez A, Holmans P, Jones L, Hardy J, Seshadri S, Schellenberg GD, Amouyel P, Williams J (2015) Common polygenic variation enhances risk prediction for Alzheimer's disease. *Brain* **138**, 3673–3684.

[4] Escott-Price V, Myers AJ, Huentelman M, Hardy J. (2017) Polygenic risk score analysis of pathologically confirmed Alzheimer disease. *Ann Neurol* **82**, 311–314.

[5] Hu Y, Li L, Ehm MG, Bing N, Song K, Nelson MR, Talmud PJ, Hingorani AD, Kumari M, Kivimäki M, Xu CF, Waterworth DM, Whittaker JC, Abecasis GR, Spino C, Kang HM. (2013) The Benefits of Using Genetic Information to Design Prevention Trials. *Am J Hum Genet* **92** (4), 547–557.

[6] Jansen I, Savage J, Watanabe K, Bryois J, Williams D, Steinberg S, Sealock J, Karlsson I, Hagg S, Athanasiu L, Voyle N, Proitsi P, Witoelar A, Stringer S, Aarsland D, Almdahl I, Andersen F, Bergh S, Bettella F, Bjornsson S, Braekhus A, Brathen G, de Leeuw C, Desikan R, Djurovic S, Dumitrescu L, Fladby T, Homan T, Jonsson P, Kiddle S, Rongve A, Saltvedt I, Sando S, Selbak G, Skene N, Snaedal J, Stordal E, Ulstein I, Wang Y, White L, Hjerling-Leffler J, Sullivan P, van der Flier W, Dobson R, Davis L, Stefansson H, Stefansson K, Pedersen N, Ripke S, Andreassen O, Posthuma D (2018) Genetic meta-analysis identifies 9 novel loci and functional pathways for Alzheimers disease risk. bioRxiv

https://www.biorxiv.org/content/early/2018/02/22/25853, Posted 22nd Feb 2018, Accessed 8th March 2018

[7] Desikan RS, Fan CC, Wang Y, Schork AJ, Cabral HJ, Cupples LA, Thompson WK, Besser L, Kukull WA, Holland D, Chen CH, Brewer JB, Karow DS, Kauppi K, Witoelar A, Karch CM, Bonham LW, Yokoyama JS, Rosen HJ, Miller BL, Dillon WP, Wilson DM, Hess CP, Pericak-Vance M, Haines JL, Farrer LA, Mayeux R, Hardy J, Goate AM, Hyman BT, Schellenberg GD, McEvoy LK, Andreassen OA, Dale AM (2013) Genetic assessment of age-associated Alzheimer disease risk: Development and validation of a polygenic hazard score. *PLoS Med* **14** (3), e1002258.

[8] Sabuncu MR, Buckner RL, Smoller JW, Lee PH, Fischl B, Sperling RA, ADNI. (2012) The association between a polygenic Alzheimer score and cortical thickness in clinically normal subjects. *Cereb Cortex* **22** (11), 2653–2661

[9] Lupton MK, Strike L, Hansell NK, Wen W, Mather KA, Armstrong NJ, Thalamuthu A, McMahon KL, de Zubicaray GI, Assareh AA, Simmons A, Proitsi P, Powell JF, Montgomery GW, Hibar DP, Westman E, Tsolaki M, Kloszewska I, Soininen H, Mecocci P, Velas B, Lovestone S, ADNI, Brodaty H, Ames D, Trollor JN, Martin NG, Thompson PM, Sachdev PS, Wright MJ (2013) The effect of increased genetic risk for Alzheimer's disease on hippocampal and amygdala volume. *Neurobiol Aging* **40**, 68–77.

[10] Voyle N, Patel H, Folarin A, Newhouse S, Johnston C, Visser PJ, Dobson RJ, Kiddle SJ, EDAR, DESCRIPA, ADNI. (2017) Genetic Risk as a Marker of Amyloid-β and Tau Burden in Cerebrospinal Fluid. *J Alzheimers Dis* **55** (4), 1417–1427.

[11] Darst BF, Koscik RL, Racine AM, Oh JM, Krause RA, Carlsson CM, Zetterberg H, Blennow K, Christian BT, Bendlin BB, Okonkwo OC, Hogan KJ, Hermann BP, Sager MA, Asthana S, Johnson SC, Engelman CD. (2017) Pathway-Specific Polygenic Risk Scores as Predictors of Amyloid-β Deposition and Cognitive Function in a Sample at Increased Risk for Alzheimer's Disease. *J Alzheimers Dis* **55** (2), 473–484.

[12] Tan CH, Fan CC, Mormino EC, Sugrue LP, Broce IJ, Hess CP, Dillion WP, Bonham LW, Yokoyama JS, Karch CM, Brewer JB, Rabinovici GD, Miller BL, Schellenberg GD, Kauppi K, Feldman HA, Holland D, McEvoy LK, Hyman BT, Andreassen OA, Dale A, Desikan RS (2017) Common polygenic variation enhances risk prediction for Alzheimer's disease. *bioRxiv* http://www.biorxiv.org/content/early/2017/07/18/165373, Posted 18 July 2017, Accessed 28 July 2017.

[13] Furney SJ, Simmons A, Breen G, Pedroso I, Lunnon K, Proitsi P, Hodges A, Powell J, Wahlund LO, Kloszewska I, Mecocci P, Soininen H, Tsolaki M, Vellas B, Spenger C, Lathrop M, Shen L, Kim S, Saykin AJ, Weiner MW, Lovestone S, ADNI, AddNeuroMed. (2011) Genome-wide association with MRI atrophy measures as a quantitative trait locus for Alzheimer's disease. *Mol Psychiatry* **16** (11), 1130–1138.

[14] Deming Y, Li Z, Kapoor M, Harari O, Del-Aguila JL, Black K, Carrell D, Cai Y, Fernandez MV, Budde J, Ma S, Saef B, Howells B, Huang KL, Bertelsen S, Fagan AM, Holtzman DM, Morris JC, Kim S, Saykin AJ, De Jager PL, Albert M, Moghekar A, O'Brien R, Riemenschneider M, Petersen RC, Blennow K, Zetterberg H, Minthon L, Van Deerlin VM, Lee VM, Shaw LM, Trojanowski JQ, Schellenberg G, Haines JL, Mayeux R, Pericak-Vance MA, Farrer LA, Peskind ER, Li G, Di Narzo AF, ADNI, ADGC, Kauwe JS, Goate AM, Cruchaga C. (2017) Genome-

wide association study identifies four novel loci associated with Alzheimer's endophenotypes and disease modifiers. *Acta Neuorpathologica* **133** (5), 839–856.

[15] Sherva R, Tripodis Y, Bennett DA, Chibnik LB, Crane PK, de Jager PL, Farrer LA, Saykin AJ, Shulman JM, Naj A, Green RC, GENAROAD, ADNI, ADGC. (2014) Genome-wide association study of the rate of cognitive decline in Alzheimer's disease. *Alzheimers Dement* **10** (1), 45–52.

[16] Allen GI, Amoroso N, Anghel C, Balagurusamy V, Bare CJ, Beaton D, Bellotti R, Bennett DA, Boehme KL, Boutros PC, Caberlotto L, Caloian C, Campbell F, Chaibub Neto E, Chang YC, Chen B, Chen CY, Chien TY, Clark T, Das S, Davatzikos C, Deng J, Dillenberger D, Dobson RJ, Dong Q, Doshi J, Duma D, Errico R, Erus G, Everett E, Fardo DW, Friend SH, Fröhlich H, Gan J, St George-Hyslop P, Ghosh SS, Glaab E, Green RC, Guan Y, Hong MY, Huang C, Hwang J, Ibrahim J, Inglese P, Iyappan A, Jiang Q, Katsumata Y, Kauwe JS, Klein A, Kong D, Krause R, Lalonde E, Lauria M, Lee E, Lin X, Liu Z, Livingstone J, Logsdon BA, Lovestone S, Ma TW, Malhotra A, Mangravite LM, Maxwell TJ, Merrill E, Nagorski J, Namasivayam A, Narayan M, Naz M, Newhouse SJ, Norman TC, Nurtdinov RN, Oyang YJ, Pawitan Y, Peng S, Peters MA, Piccolo SR, Praveen P, Priami C, Sabelnykova VY, Senger P, Shen X, Simmons A, Sotiras A, Stolovitzky G, Tangaro S, Tateo A, Tung YA, Tustison NJ, Varol E, Vradenburg G, Weiner MW, Xiao G, Xie L, Xie Y, Xu J, Yang H, Zhan X, Zhou Y, Zhu F, Zhu H, Zhu S, ADNI. (2016) Crowdsourced estimation of cognitive decline and resilience in Alzheimer's disease. *Alzheimers Dement* **12** (6), 645–653

[17] Olsson B, Lautner R, Andreasson U, Öhrfelt A, Portelius E, Bjerke M, Hölttä M, Rosén C, Olsson C, Strobel G, Wu E, Dakin K, Petzold M, Blennow K, Zetterberg H. (2016) CSF and

blood biomarkers for the diagnosis of Alzheimer's disease: a systematic review and meta-analysis. *Lancet Neurol* **15** (7), 673–684.

[18] Chiu M, Yang S, Horng H, Yang C, Chen T, Ho C. (2013) Combined plasma biomarkers for diagnosing mild cognition impairment and Alzheimer's disease. *ACS Chem Neurosci* **4** (12), 1530–1536.

[19] Kiddle SJ, Sattlecker M, Proitsi P, Simmons A, Westman E, Bazenet C, Nelson SK, Williams S, Hodges A, Johnston C, Soininen H, Kłoszewska I, Mecocci P, Tsolaki M, Vellas B, Newhouse S, Lovestone S, Dobson RJ. (2014) Candidate blood proteome markers of Alzheimer's disease onset and progression: a systematic review and replication study. *J Alzheimers Dis* **38** (3), 515–531.

[20] Baird A, Westwood S, Lovestone S. (2015) Blood-Based Proteomic Biomarkers of Alzheimer's Disease Pathology. *Front Neurol* **6**, 236.

[21] Nakamura A, Kaneko N, Villemagne VL, Kato T, Doecke J, Doré V, Fowler C, Li Q-X, Martins R, Rowe C, Tomita T, Matsuzaki K, Ishii K, Ishii K, Arahata Y, Iwamoto S, Ito K, Tanaka K, Masters CL, Yanagisawa K (2018) High performance plasma amyloid-β biomarkers for Alzheimer's disease. *Nature* **554**, 249-254.

[22] Mapstone M, Cheema AK, Fiandaca MS, Zhong X, Mhyre TR, MacArthur LH, Hall WJ, Fisher SG, Peterson DR, Haley JM, Nazar MD, Rich SA, Berlau DJ, Peltz CB, Tan MT, Kawas CH, Federoff HJ. (2014) Plasma phospholipids identify antecedent memory impairment in older adults. *Nat Med* **20**, 415–418.

[23] Proitsi P, Kim M, Whiley L, Simmons A, Sattlecker M, Velayudhan L, Lupton MK, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Lovestone S, Powell JF, Dobson RJ,

Legido-Quigley C. (2017) Association of blood lipids with Alzheimer's disease: A comprehensive lipidomics analysis. *Alzheimers Dement* **13** (2), 140–151.

[24] Han G, Wang J, Zeng F, Feng X, Yu J, Cao HY, Yi X, Zhou H, Jin LW, Duan Y, Wang YJ, Lei H. (2013) Characteristic transformation of blood transcriptome in Alzheimer's disease. *J Alzheimers Dis* **35** (2), 373–386.

[25] Voyle N, Keohane A, Newhouse S, Lunnon K, Johnston C, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Lovestone S, Hodges A, Kiddle S, Dobson RJ. (2016) A Pathway Based Classification Method for Analyzing Gene Expression for Alzheimer's Disease Diagnosis. *J Alzheimers Dis* **49** (3), 659–669.

[26] Lunnon K, Sattlecker M, Furney SJ, Coppola G, Simmons A, Proitsi P, Lupton MK, Lourdusamy A, Johnston C, Soininen H, Kłoszewska I, Mecocci P, Tsolaki M, Vellas B, Geschwind D, Lovestone S, Dobson R, Hodges A, AddNeuroMed. (2012) A blood gene expression marker of early Alzheimer's disease. *J Alzheimers Dis* **33** (3), 737–753.

[27] Voyle N, Kim M, Proitsi P, Ashton NJ, Baird AL, Bazenet C, Hye A, Westwood S, Chung R, Ward M, Rabinovici GD, Lovestone S, Breen G, Legido-Quigley C, Dobson RJ, Kiddle SJ, ADNI. (2016) Blood metabolite markers of neocortical amyloid-β burden: discovery and enrichment using candidate proteins. *Trans Psychiatry* **6**, e719.

[28] Choi J, Malakowsky CA, Talent JM, Conrad CC, Gracy RW. (2002) Identification of oxidized plasma proteins in Alzheimer's disease. *J Biochem Biophys Res Commun* **293** (5), 566–570.

[29] Doecke JD, Laws SM, Faux NG, Wilson W, Burnham SC, Lam CP, Mondal A, Bedo J, Bush AI, Brown B, De Ruyck K, Ellis KA, Fowler C, Gupta VB, Head R, Macaulay SL, Pertile K, Rowe CC, Rembach A, Rodrigues M, Rumble R, Szoeke C, Taddei K, Taddei T, Trounson B, Ames D, Masters CL, Martins RN, ADNI, AIBL. (2012) Blood-based protein biomarkers for diagnosis of Alzheimer disease. *Arch Neurol* **293** (5), 1318–1325.

[30] Kiddle SJ, Thambisetty M, Simmons A, Riddoch-Contreras J, Hye A, Westman E, Pike I, Ward M, Johnston C, Lupton MK, Lunnon K, Soininen H, Kloszewska I, Tsolaki M, Vellas B, Mecocci P, Lovestone S, Newhouse S, Dobson R, ADNI (2012) Plasma based markers of [11C] PiB-PET brain amyloid burden. *J Alzheimers Dis* **7**, e44260.

[31] Henkel AW, Müller K, Lewczuk P, Müller T, Marcus K, Kornhuber J, Wiltfang J. (2012) Multidimensional plasma protein separation technique for identification of potential Alzheimer's disease plasma biomarkers: a pilot study. *J Neural Transm (Vienna)* **119**, 779–788.

[32] Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. (2001) Replication validity of genetic association studies. *Nat Genet* **29**, 306–309.

[33] Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, Matthews P, Ong G, Pell J, Silman A, Young A, Sprosen T, Peakman T, Collins R. (2015) UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* **12** (3), e1001779.

[34] Collins F. (2015) A new iniative on precision medicine. *N Engl J Med* **372**, 793–795.

[35] O'Bryant SE, Mielke MM, Rissman RA, Lista S, Vanderstichele H, Zetterberg H, Lewczuk P, Posner H, Hall J, Johnson L, Fong YL, Luthman J, Jeromin A, Batrla-Utermann R, Villarreal

A, Britton G, Snyder PJ, Henriksen K, Grammas P, Gupta V, Martins R, Hampel H; BBB PIA. (2017) Blood-based biomarkers in Alzheimer disease: Current state of the science and a novel collaborative paradigm for advancing from discovery to clinic. *Alzheimers Dement* **13** (1), 45–58.

[36] Kiddle SJ, Steves CJ, Mehta M, Simmons A, Xu X, Newhouse S, Sattlecker M, Ashton NJ, Bazenet C, Killick R, Adnan J, Westman E, Nelson S, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Curtis C, Breen G, Williams SC, Lovestone S, Spector TD, Dobson RJ. (2015) Plasma protein biomarkers of Alzheimer's disease endophenotypes in asymptomatic older twins: early cognitive decline and regional brain volumes. *Trans Psychiatry* **5**, e584.

[37] Thambisetty M, Tripaldi R, Riddoch-Contreras J, Hye A, An Y, Campbell J, Sojkova J, Kinsey A, Lynham S, Zhou Y, Ferrucci L, Wong DF, Lovestone S, Resnick SM. (2010) Proteome-based plasma markers of brain amyloid-β deposition in non-demented older individuals. *J Alzheimers Dis* **22** (4), 1099–1109.

[38] Westwood S, Leoni E, Hye A, Lynham S, Khondoker MR, Ashton NJ, Kiddle SJ, Baird AL, Sainz-Fuertes R, Leung R, Graf J, Hehir CT, Baker D, Cereda C, Bazenet C, Ward M, Thambisetty M, Lovestone S. (2016) Blood-Based Biomarker Candidates of Cerebral Amyloid Using PiB PET in Non-Demented Elderly. *J Alzheimers Dis* **52** (2), 561–572.

[39] Ijsselstijn L, Dekker LJ, Stingl C, van der Weiden MM, Hofman A, Kros JM, Koudstaal PJ, Sillevis Smitt PA, Ikram MA, Breteler MM, Luider TM. (2011) Serum levels of pregnancy zone protein are elevated in presymptomatic Alzheimer's disease. *J Proteome Res* **10** (11), 4902-4910.

[40] Loong TW. (2003) Understanding sensitivity and specificity with the right side of the brain. *BMJ* **327**, 716-719.

[41] Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, Woodward M. (2012) Risk prediction models: II. External validation, model updating, and impact assessment. *Heart* **98** (9), 691–698.

[42] Kohavi R. (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection in *IJCAI'95 Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2*, Stanford, CA, 1137 – 1145.

[43] Marksteiner J, Kemmler G, Weiss EM, Knaus G, Ullrich C, Mechtcheriakov S, Oberbauer H, Auffinger S, Hinterhölzl J, Hinterhuber H, Humpel C. (2011) Five out of 16 plasma signaling proteins are enhanced in plasma of patients with mild cognitive impairment and Alzheimer's disease. *Neurobiol Aging* **32** (3), 539–540.

[44] Li D, Misialek JR, Boerwinkle E, Gottesman RF, Sharrett AR, Mosley TH, Coresh J, Wruck LM, Knopman DS, Alonso A. (2016) Plasma phospholipids and prevalence of mild cognitive impairment and/or dementia in the ARIC Neurocognitive Study (ARIC-NCS). *Alzheimers Dement (Amst)* **3**, 73–82.

[45] Casanova R, Varma S, Simpson B, Kim M, An Y, Saldana S, Riveros C, Moscato P, Griswold M, Sonntag D, Wahrheit J, Klavins K, Jonsson PV, Eiriksdottir G, Aspelund T, Launer LJ, Gudnason V, Legido Quigley C, Thambisetty M. (2016) Blood metabolite markers of preclinical Alzheimer's disease in two longitudinally followed cohorts of older individuals. *Alzheimers Dement* **12** (7), 815-822.

[46] Ray S, Britschgi M, Herbert C, Takeda-Uchimura Y, Boxer A, Blennow K, Friedman LF, Galasko DR, Jutel M, Karydas A, Kaye JA, Leszek J, Miller BL, Minthon L, Quinn JF, Rabinovici GD, Robinson WH, Sabbagh MN, So YT, Sparks DL, Tabaton M, Tinklenberg J,

Yesavage JA, Tibshirani R, Wyss-Coray T. (2007) Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins. *Nat Med* **13** (11), 1359–1362.

[47] Wason P. (1960) On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology* **12**, 129–140.

[48] Kerr N. (1998) Hypothesizing After Results are Known. *Personality and Social Psychology Bulletin* **2** (3), 196-217.

[49] Rosenthal R. (1979) The file drawer problem and tolerance for null results. *Psychological Bulletin* **86** (3), 638–641.

[50] Ioannidis J. (2005) Why most published research findings are false. *PLoS Med* **2** (8), e124.

[51] Voyle N, Baker D, Burnham SC, Covin A, Zhang Z, Sangurdekar DP, Tan Hehir CA, Bazenet C, Lovestone S, Kiddle S, Dobson RJ, AIBL. (2015) Blood Protein Markers of Neocortical Amyloid-β Burden: A Candidate Study Using SOMAscan Technology. *J Alzheimers Dis* **46** (4), 947-961.

[52] Ioannidis J. (2005) Why most discovered true associations are inflated. *Epidemiology* **19** (5), 640-648.

[53] Menni C, Kiddle SJ, Mangino M, Viñuela A, Psatha M, Steves C, Sattlecker M, Buil A, Newhouse S, Nelson S, Williams S, Voyle N, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Lovestone S, Spector TD, Dobson R, Valdes AM. (2015) Circulating Proteomic Signatures of Chronological Age. *J Gerontol A Biol Sci Med Sci* **70** (7), 809–816.

[54] Di Narzo AF, Telesco SE, Brodmerkel C, Argmann C, Peters LA, Li K, Kidd B, Dudley J, Cho J, Schadt EE, Kasarskis A, Dobrin R, Hao K. (2017) High-Throughput Characterization of

Blood Serum Proteomics of IBD Patients with Respect to Aging and Genetic Factors. *PLoS Genet* **13** (1), e1006565.

[55] Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, Burgess S, Jiang T, Paige E, Surendran P, Oliver-Williams C, Kamat MA, Prins BP, Wilcox SK, Zimmerman ES, Chi A, Bansal N, Spain SL, Wood AM, Morrell NW, Bradley JR, Janjic N, Roberts DJ, Ouwehand WH, Todd JA, Soranzo N, Suhre K, Paul DS, Fox CS, Plenge RM, Danesh J, Runz H, Butterworth AS. (2017) Consequences Of Natural Perturbations In The Human Plasma Proteome. *bioRxiv* http://www.biorxiv.org/content/early/2017/05/05/134551, Posted 5 May 2017, Accessed 28 July 2017.

[56] Ioannidis J. (2014) How to make more published research true. *PLoS Med* **11** (10), e1001747.

[57] Collins G, Reitsma J, Altman D, Moons K. (2015) Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med* **162** (1), 55-63.

[58] O'Bryant SE, Gupta V, Henriksen K, Edwards M, Jeromin A, Lista S, Bazenet C, Soares H, Lovestone S, Hampel H, Montine T, Blennow K, Foroud T, Carrillo M, Graff-Radford N, Laske C, Breteler M, Shaw L, Trojanowski JQ, Schupf N, Rissman RA, Fagan AM, Oberoi P, Umek R, Weiner MW, Grammas P, Posner H, Martins R, STAR-B, BBBIG. (2015) Guidelines for the standardization of preanalytic variables for blood-based biomarker studies in Alzheimer's disease research. *Alzheimers Dement* **11** (5), 549–560.

[59] Longo D. (2016) Data sharing. *N Engl J Med* **374**, 276–277.

[60] Kiddle S, Parodi A, Johnston C, Wallace C, Dobson R. (2016) Heterogeneity of cognitive decline in dementia: a failed attempt to take into account variable time-zero severity. *bioRxiv* http://www.biorxiv.org/content/early/2017/07/11/060830, Posted 11 July 2016, Accessed 28 July 2017.

[61] Perera G, Broadbent M, Callard F, Chang CK, Downs J, Dutta R, Fernandes A, Hayes RD, Henderson M, Jackson R, Jewell A, Kadra G, Little R, Pritchard M, Shetty H, Tulloch A, Stewart R. (2016) Cohort profile of the South London and Maudsley NHS Foundation Trust Biomedical Research Centre (SLaM BRC) Case Register: current status and recent enhancement of an Electronic Mental Health Record-derived data resource. *BMJ Open* **6** (3), e008721.

[62] Jackson R, Agrawal A, Lui K, Folarin A, Wu H, Groza T, Roberts A, Gorrell G, Song X, Lewsley D, Northwood D, Stringer C, Stewart R, Dobson R. (2016) CogStack - Experiences of Deploying Integrated Information Retrieval And Extraction Services In A Large National Health Service Foundation Trust Hospital. *bioRxiv* http://www.biorxiv.org/content/early/2017/04/02/123299, Posted 2 April 2017, Accessed 28 July 2017.

[63] Baker E, Iqbal E, Johnston C, Broadbent M, Shetty H, Stewart R, Howard R, Newhouse S, Khondoker M, Dobson RJB. (2017) Trajectories of dementia-related cognitive decline in a large mental health records derived patient cohort. *PLoS One* **12** (6), e0178562.