

Evaluation of a Previously Suggested Plasma Biomarker Panel to Identify Alzheimer's Disease

Maria Björkqvist^{1*}, Mattias Ohlsson², Lennart Minthon^{3,4}, Oskar Hansson^{3,4}

1 Brain Disease Biomarker Unit, Department of Experimental Medical Science, Wallenberg Neuroscience Center, Lund University, Lund, Sweden, **2** Computational Biology and Biological Physics, Lund University, Lund, Sweden, **3** Clinical Memory Research Unit, Department of Clinical Sciences Malmö, Lund University, Malmö, Sweden, **4** Neuropsychiatric Clinic, Skåne University Hospital, Malmö, Sweden

Abstract

There is an urgent need for biomarkers in plasma to identify Alzheimer's disease (AD). It has previously been shown that a signature of 18 plasma proteins can identify AD during pre-dementia and dementia stages (Ray et al, *Nature Medicine*, 2007). We quantified the same 18 proteins in plasma from 174 controls, 142 patients with AD, and 88 patients with other dementias. Only three of these proteins (EGF, PDG-BB and MIP-1 δ) differed significantly in plasma between controls and AD. The 18 proteins could classify patients with AD from controls with low diagnostic precision (area under the ROC curve was 63%). Moreover, they could not distinguish AD from other dementias. In conclusion, independent validation of results is important in explorative biomarker studies.

Citation: Björkqvist M, Ohlsson M, Minthon L, Hansson O (2012) Evaluation of a Previously Suggested Plasma Biomarker Panel to Identify Alzheimer's Disease. *PLoS ONE* 7(1): e29868. doi:10.1371/journal.pone.0029868

Editor: Michelle L. Block, Virginia Commonwealth University, United States of America

Received: July 11, 2011; **Accepted:** December 5, 2011; **Published:** January 18, 2012

Copyright: © 2012 Björkqvist et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the Swedish Research Council, the Swedish Brain Power, the Regional Agreement on Medical Training and Clinical Research (ALF) between Skåne County Council and Lund University and the Torsten and Ragnar Söderberg Foundation. MB is supported by BenteRexed foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: maria.bjorkqvist@med.lu.se

Introduction

Alzheimer's disease (AD) is the major cause of dementia and a great medical and socioeconomic problem worldwide. As populations get older, the prevalence of AD will increase considerably during the coming decades [1]. The pathological characteristics of AD are senile plaques and neurofibrillary tangles, containing aggregated amyloid β (A β) and hyperphosphorylated tau protein, respectively [1,2]. A β accumulation is thought to start many decades before symptoms occur [3]. During the last few years, it has become more apparent that disease-modifying therapies for AD are more likely to be successful if initiated during the early stages of the disease when neurodegeneration is not yet too severe [4,5]. Therefore, biomarkers are urgently needed to correctly identify subjects affected by AD before they have developed dementia [5,6]. Cerebrospinal fluid biomarker can identify prodromal AD with acceptable accuracy [7–9]. However, plasma is much easier obtained than cerebrospinal fluid. Therefore, it was a major breakthrough when Ray and collaborators found that a pattern of 18 proteins in plasma could classify samples from AD and controls with almost 90% accuracy [10]. The same plasma proteins could also predict the patients with mild cognitive impairment who would later develop AD. The study comprised of 259 plasma samples obtained from in total 7 clinical centres [10].

In the present study, we evaluated the diagnostic value of the same 18 proteins as Ray et al [10], using 433 plasma samples obtained at Skåne University Hospital, Sweden, from 174 controls, 142 patients with AD, 29 patients with depression, and 88 patients with other types of dementia than AD (i.e 37 with Lewy Body dementia, 11 with Parkinson's disease with dementia, 22 with frontotemporal dementia, 18 with vascular dementia).

Materials and Methods

Collection and processing of human plasma samples

The study population was recruited at the memory disorder clinic, Skåne University Hospital, Malmö, Sweden. The patients underwent thorough standard examinations conducted by a trained physician, including neurological, physical and psychiatric examinations. The patients who during clinical follow-up received a diagnosis of AD had to meet the DSM-III-R criteria of dementia [11] and the criteria of probable AD defined by NINCDS-ADRDA [12]. Subjects who were diagnosed as having vascular dementia (VaD) fulfilled the DSM-III-R criteria of dementia and the requirements for probable VaD by NINDS-AIREN [13] or the recommendations by Erkinjuntti and co-workers for VaD of the subcortical type [14]. For patients who developed dementia with Lewy bodies (DLB) or frontotemporal dementia, the consensus criteria by McKeith and collaborators [15] and McKhann and colleagues were used [16], respectively. The healthy volunteers had no memory complaints or other cognitive symptoms, preservation of general cognitive function, and no active neurological or psychiatric diseases.

The study was conducted in accordance with the Helsinki Declaration and approved by the ethics committee of Lund University, Sweden. All subjects gave informed written consent.

Non-fasting plasma was collected between 9 and 11 am. After venipuncture, blood was collected in tubes prepared with EDTA to prevent coagulation. Samples were centrifuged, and plasma was removed from the tubes leaving 1 ml of plasma to avoid contamination of plasma with blood cells including trombocytes. Within one hour from venipuncture the plasma was frozen in polypropylene tubes at -80°C until biochemical analysis.

Analysis of plasma proteins

Quantibody[®] Human Costum Cytokine Antibody Array was performed by RayBiotech (as per company description) on blinded samples for the following markers; ANG-2I, CAM-1, IGFBP-6, PARC, PDGF-BB, RANTES, EGF, G-CSF, GDNF, IL-1a, IL-3, IL-8, IL-11, MCP-3, M-CSF, MIP-1δ, TNFa, and TRAIL R4. A positive control (four biotin-labelled bovine IgG spot) was included on each array and used for inter- and intra-slide normalization. For a quadruplicate spot, outliers as value above 30% over the median, was excluded. All samples were analyzed in a single run to minimize variation.

Selected cytokines (M-CSF and TNF- α) were quantified in triplicates using Meso Scale Discovery (MSD[®], Gaithersburg, MD) electrochemoluminescence assays using a modification of the manufacturer's protocol. 30 μ l was used as the sample volume and a 10-point standard curve was used, ranging from 2500 pg/ml to 0 pg/ml. The sample and calibrator were incubated on the MSD plate for 3 h (instead of 2 h), followed by a wash (as per manufacturer's recommendation). The MSD plate was then incubated with detection antibody solution for 3 h (instead of 2 h) before wash and read as per manufacturer's recommendation. Results were analyzed on a SECTORTM 6000 instrument (MSD).

The operator was unaware of the disease state of each sample during processing and statistical analysis was performed independently.

Statistical analysis

The statistical analyses were accomplished using SPSS for Windows, version 18.0.1 (SPSS Inc/IBM, Chicago, IL, USA). To compare demographic and plasma data between groups, non-parametric Kruskal-Wallis tests were performed followed by Mann-Whitney *U*-tests for continuous variables. Pearson's χ^2 test was used for dichotomous variables.

To assess the ability of the plasma data to separate groups (AD vs. Controls or AD vs. other dementia) multiple logistic regression [17], artificial neural network (ANN) [18] and nearest shrunken centroid [19] classification models were used. The latter was the method used by Ray et al. Bagging ensembles [20] of standard multi-layer perceptrons with one hidden layer were used in the ANN models. The size of the ensemble was set to 30 and the number of hidden nodes of the individual networks in the ensemble was two. No effort was made to tune these parameters. The nearest shrunken centroid method was implemented using the R package *panvr*. The area under the ROC curve (AUC) was used

Table 1. Subject demographics and plasma protein levels.

	Controls (n = 174)	Depr (n = 29)	AD (n = 142)	FTD (n = 22)	VaD (n = 18)	PDD (n = 11)	DLB (n = 37)
Mean age (range)	74 (62–99)	59 (42–76)	76 ^{a,b} (56–87)	62 (43–78)	76 (56–84)	72 (62–81)	74 (54–85)
Women	117	15	40 ^c	11	14	5	26
MMSE	29 \pm 0,1	28 \pm 0,4	21 \pm 0,4 ^{d,e}	22 \pm 1,1	22 \pm 0,8	20 \pm 2,0	21 \pm 0,9
ANG-2	2444 \pm 143	1710 \pm 232	2306 \pm 284	2043 \pm 320	2410 \pm 527	1789 \pm 535	2335 \pm 341
ICAM-1	224 \pm 9	250 \pm 31	216 \pm 9	222 \pm 28	220 \pm 22	162 \pm 16	204 \pm 15
IGFBP-6	382 \pm 13	366 \pm 21	377 \pm 13	337 \pm 19	373 \pm 40	391 \pm 51	379 \pm 17
PARC (CCL18)	43 \pm 2,7	40 \pm 6,0	40 \pm 2,4	38 \pm 5,1	39 \pm 7,9	33 \pm 5,3	41 \pm 3,9
PDGF-BB	4385 \pm 288	6339 \pm 1122	5701 \pm 427 ^f	5896 \pm 706	9302 \pm 3278	8557 \pm 3513	5798 \pm 1431
RANTES (CCL5)	18 \pm 0,7	20 \pm 1,5	20 \pm 0,7	20 \pm 1,3	23 \pm 3,2	19 \pm 2,3	18 \pm 1,3
EGF	541 \pm 31	665 \pm 65	706 \pm 32 ^g	825 \pm 79	799 \pm 158	907 \pm 117	580 \pm 49
G-CSF	54 \pm 3,1	41 \pm 5,1	57 \pm 6	56 \pm 10	53 \pm 12	56 \pm 13	42 \pm 3,9
GDNF	103 \pm 33	86 \pm 34	322 \pm 262	79 \pm 15	41 \pm 11	34 \pm 12	45 \pm 6,5
IL-1a	13 \pm 2,4	11,2 \pm 3,2	26 \pm 17	11 \pm 1,7	8,5 \pm 1,9	10 \pm 2,0	9,5 \pm 1,3
IL-3	44 \pm 15	29 \pm 9,1	87 \pm 63	32 \pm 6,3	14 \pm 3,8	15 \pm 4,1	22 \pm 6,0
IL-8 (CXCL8)	11 \pm 1,0	8,5 \pm 0,8	10 \pm 0,5	12 \pm 1,9	10 \pm 1,4	11 \pm 1,2	10 \pm 0,8
IL-11	236 \pm 27	185 \pm 78	249 \pm 34	249 \pm 53	260 \pm 74	158 \pm 563	263 \pm 49
MCP-3 (CCL7)	51 \pm 4,8	47 \pm 6,4	82 \pm 30	51 \pm 6,8	52 \pm 12	67 \pm 8,2	42 \pm 6,1
M-CSF	1,0 \pm 0,2	1,3 \pm 0,7	1 \pm 0,3	1,2 \pm 0,37	0,77 \pm 0,51	1,2 \pm 0,5	1,0 \pm 0,4
MIP-1d (CCL15)	2361 \pm 85	2322 \pm 229	2845 \pm 142 ^h	2452 \pm 200	2383 \pm 312	3154 \pm 453	3144 \pm 348
TNF-a	23 \pm 4,6	15 \pm 2,5	17 \pm 1,3	17 \pm 2,6	15 \pm 2,8	16 \pm 2,6	16 \pm 2,0
TRAILR4	36 \pm 15	12 \pm 4,3	169 \pm 157	17 \pm 7,1	13 \pm 9,1	3,1 \pm 2,3	9,5 \pm 3,3

Table 1. Values are given in ng/L, except for ICAM-1, IGFBP-6, PARC and RANTES for which concentrations are given in μ g/L. When comparing AD vs 1) Controls, 2) Depression and 3) other dementias the following significant changes were observed:

^aAD vs control $p=0.01$,

^bAD vs depression $p<0.001$, AD vs other dementias $p=0.01$.

^cAD vs depression $p=0.034$.

^dAD vs control $p<0.001$,

^eAD vs depression $p=0.001$.

^fAD vs control $p=0.004$.

^gAD vs control $p<0.001$.

^hAD vs control $p=0.011$ (Kruskal–Wallis one-way analysis of variance by ranks followed by Mann-Whitney *U* test).

Abbreviations: Depr, depression; AD, Alzheimers disease; FTD, frontotemporal lobe dementia; VaD, vascular dementia; PDD, Parkinson's disease with dementia; DLB = dementia with Lewy bodies.

doi:10.1371/journal.pone.0029868.t001

to measure the performance of the classification models. For all three models, 10-fold cross-validation was used to estimate true AUC values. The cross-validation procedure was repeated 100 times, with random 10-fold splits each time, in order to decrease random fluctuations.

Results

In table 1 we present the demographic data and levels of the 18 plasma proteins obtained when using Quantibody® Human Costum Cytokine Antibody Array (RayBiotech). The subjects affected by AD were slightly older than the controls and the group affected by other forms of dementias ($p \leq 0.01$). Only three proteins of the 18 proteins, (EGF, PDG-BB and MIP-1 δ), were found to be significantly altered in plasma from AD patients when compared to controls (table 1). None of the proteins differed between the AD group and the group with other dementias than AD. Analyses of two cytokines (M-CSF and TNF- α) with ELISA technology verified that there were no statistical differences between AD and control plasma samples (in control plasma, $n = 148$, M-CSF levels were 21.82 ± 0.87 ng/L and TNF- α levels were 1.90 ± 0.14 ng/L, in AD plasma, $n = 148$, the corresponding levels were 24.03 ± 0.71 and 1.85 ± 0.06 ng/L respectively).

When classifying the AD group from the controls, the cross-validation AUC for the logistic regression model was 0.60 using all 18 proteins. The corresponding AUC for the ANN model and the nearest shrunken centroid classifier was 0.63. When only using the three proteins that differed significantly between groups (i.e. EGF, PDG-BB and MIP-1 δ), as inputs to the classifiers, the AUC increased to 0.66 for all three models.

A worse performance was obtained when classifying the group with AD from the group with other forms of dementia than AD. Using all plasma proteins the cross-validation AUC was below 0.5 indicating no classification ability at all. This was true for all three models. The best individual protein in terms of AUC performance was TRAIL-R4 with an AUC of 0.61 (cross-validation result).

To further illustrate the limitation of the 18 plasma protein panel to differentiate AD from the controls and other dementia groups, multidimensional scaling (MDS) plots were produced (figure 1). These plots show a large degree of overlap between the diagnostic groups.

Discussion

Characterizing protein markers in plasma has created optimism for finding detectable disease-specific pattern of changes. A biomarker panel of eighteen plasma proteins were shown in 2007 to classify blinded samples from AD and control subjects with close to 90% accuracy and to identify patients who had mild cognitive impairment that progressed to Alzheimer's disease [10]. The study was comprised of 259 plasma samples obtained from in total seven different clinical centres.

Interestingly, when re-analysing the same data set, originally obtained from the Ray et al study, a subset of plasma proteins (as z-scores of plasma proteins) resulted in good diagnostic accuracy [21]. However, following up on these results, using bead-based multiplex technology, Soares and co-workers have shown that when using a subset of the proteins included in the original 18 protein panel a diagnostic accuracy of only 61% was obtained when differentiating cases with AD from controls [22]. Later Rocha de Paula et al. proposed, using the original data set provided by Ray et al, that including pair-wise differences of z-score values to the mathematical method, could collectively provide a good discrimination value [23].

In the present study we found that the 18 plasma protein panel could classify samples from AD and controls with an AUC of only 63%, indicating that this protein panel cannot be used in the clinical diagnostic work-up of AD. The same protein panel could not distinguish cases with AD from subjects affected by other forms of dementia. In addition, the pattern of protein changes observed in the present study was not the same as in Ray et al. More specifically, in the training set described in the study by Ray

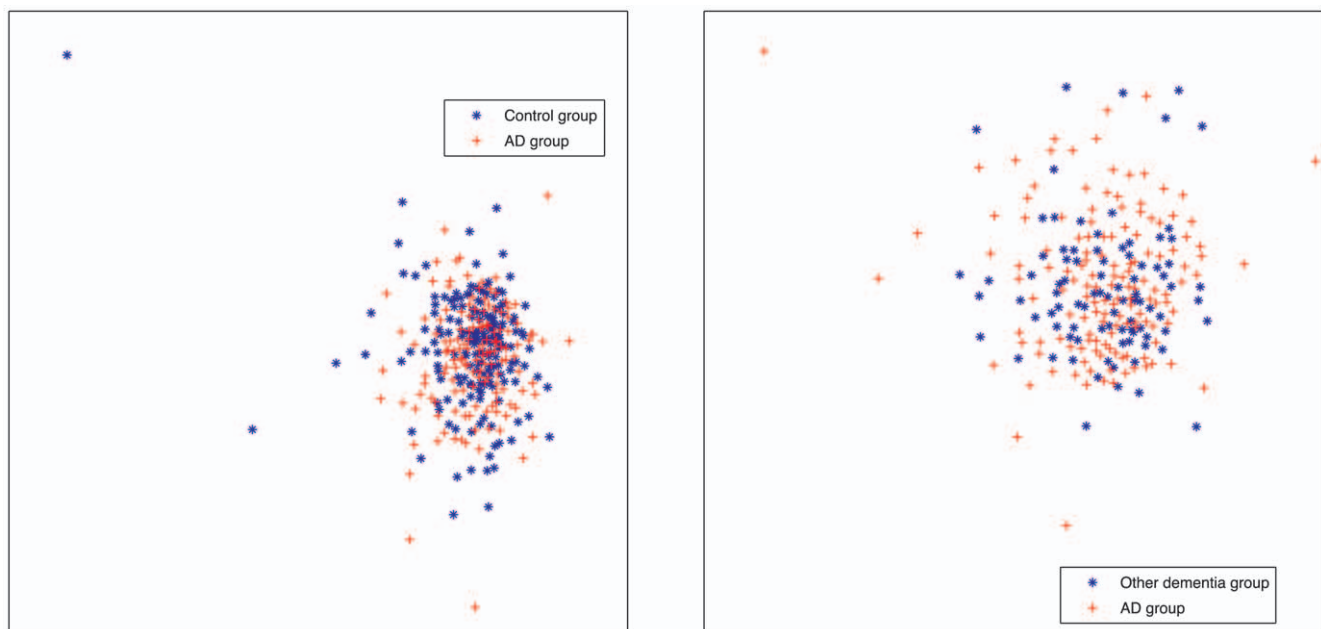


Figure 1. Multidimensional scaling plots (MDS) for the data. The left figure shows a MDS projection to 2 dimensions, using all 18 proteins, for the AD and the Control groups. The right figure is the corresponding plot for the AD and the other dementias group.
doi:10.1371/journal.pone.0029868.g001

et al [10], plasma α -levels of PDGF-BB, EGF and MIP-1 δ were seen to be reduced in AD cases. In contradiction to this, in the present study the plasma levels of these proteins were increased in the AD cases. Similarly, Marksteiner et al have found that plasma MIP-1 δ and EGF are increased in AD patients when compared to patients affected by depression [24].

Several of the 18 proteins included in the biomarker panel are involved in the immune response [10]. There are, however, important caveats to the use of plasma immune markers as biomarkers of disease progression or diagnostic predictors. AD is a slowly progressive disorder and systemic changes in the blood are likely to be subtle and difficult to monitor. There are also technical limitations in assaying low abundant cytokines and many factors likely influence plasma immune markers, such as concomitant infection and inflammatory illness. Furthermore, many cytokines has been shown to display diurnal variation [25] and different handlings as well as storage of samples are known to affect the levels of many biomarkers. Therefore, standardization of pre-analytical procedures is vital to obtain reproducible results. To increase the possibility of successful reproduction of biomarker studies in the future the handling of samples should be carefully described, including data describing the time from venipuncture to minus 80 freezer storage, time of day that venipuncture was performed and if samples were collected fasting or non-fasting. Moreover, when using samples from different clinical centres all diagnostic groups (including controls) need to be obtained from

each clinical centre in order to be able to investigate potential variations in biomarker levels between different clinical sites.

A limitation of this study is that the array-based method used is a potentially unreliable tool to disprove the original study by Ray et al, who also used a similar and non-validated method. However, in the present study we selected two cytokines for confirmation measurements with a standardized ELISA method, and again found no statistical differences between control and AD samples.

Our results indicate that multiplex platforms might be important for biomarker discovery, but validation of the results using new patient cohorts as well as other analytical techniques are vital. At least two patient-control cohorts with all important diagnostic groups present in each will likely be needed to verify obtained data. Importantly, recent data show that highly cited biomarker studies often report larger effect estimates than are reported in subsequent meta analyses [26]. This further strengthens our notion that validation is crucial in biomarker research.

Acknowledgments

We thank Professor Carsten Peterson for valuable discussions.

Author Contributions

Conceived and designed the experiments: MB OH MO. Performed the experiments: MB MO. Analyzed the data: MB MO OH. Contributed reagents/materials/analysis tools: MB OH MO. Wrote the paper: MB OH MO LM.

References

- Blennow K, de Leon MJ, Zetterberg H (2006) Alzheimer's disease. *Lancet* 368: 387–403.
- Querfurth HW, LaFerla FM (2010) Alzheimer's disease. *N Engl J Med* 362: 329–344.
- Jack CRJ, Knopman DS, Jagust WJ, Shaw LM, Aisen PS, et al. (2010) Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurology* 9: 119–128.
- Perrin RJ, Fagan AM, Holtzman DM (2009) Multimodal techniques for diagnosis and prognosis of Alzheimer's disease. *Nature* 461: 916–922.
- Blennow K, Hampel H, Weiner M, Zetterberg H (2010) Cerebrospinal fluid and plasma biomarkers in Alzheimer disease. *Nat Rev Neurol* 6: 131–144.
- Hampel H, Frank R, Broich K, Teipel SJ, Katz RG, et al. (2010) Biomarkers for Alzheimer's disease: academic, industry and regulatory perspectives. *Nat Rev Drug Discov* 9: 560–574.
- Mattsson N, Zetterberg H, Hansson O, Andreassen N, Parnetti L, et al. (2009) CSF biomarkers and incipient Alzheimer disease in patients with mild cognitive impairment. *JAMA* 302: 385–393.
- Hansson O, Zetterberg H, Buchhave P, Londos E, Blennow K, et al. (2006) Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impairment: a follow-up study. *Lancet Neurol* 5: 228–234.
- Shaw LM, Vanderstichele H, Knopik-Czajka M, Clark CM, Aisen PS, et al. (2009) Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. *Ann Neurol* 65: 403–413.
- Ray S, Britschgi M, Herbert C, Takeda-Ushimara, Boxer A, et al. (2007) Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins. *Nat Med* 13: 1359–1362.
- Association AP (1987) *Diagnostic and Statistical Manual of Mental Disorders*, third, revised ed. Washington DC: American Psychiatric Association.
- McKhann G, Drachman D, Folstein M, Folstein M, Katzman R, et al. (1984) Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 34: 939–944.
- Roman GC, Tatemichi TK, Erkinjuntti T, Cummings JL, Masde JC, et al. (1993) Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology* 43: 250–260.
- Erkinjuntti T, Inzitari D, Pantoni L, Wallin A, Scheltens P, et al. (2000) Research criteria for subcortical vascular dementia in clinical trials. *J Neural Transm Suppl* 59: 23–30.
- McKeith IG, Perry EK, Perry RH (1999) Report of the second dementia with Lewy body international workshop: diagnosis and treatment. Consortium on Dementia with Lewy Bodies. *Neurology* 53: 902–905.
- McKhann GM, Albert MS, Grossman M, Miller B, Dickson D, et al. (2001) Clinical and pathological diagnosis of frontotemporal dementia: report of the Work Group on Frontotemporal Dementia and Pick's Disease. *Arch Neurol* 58: 1803–1809.
- Hosmer D, Lemeshow S (1989) *Applied logistic regression*. New York: Wiley.
- Bishop CM (1995) *Neural Networks for Pattern Recognition*. Oxford: Oxford University Press.
- Tibshirani R, Hastie T, Narasimhan B, Chu G (2002) Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proceedings of the National Academy of Sciences of the United States of America* 99: 6567–72.
- Breiman L (1996) Bagging predictors. *Machine Learning* 24: 123–140.
- Gómez Ravetti M, Moscato P (2008) Identification of a 5-protein biomarker molecular signature for predicting Alzheimer's disease. *PLoS ONE* 3: e3111.
- Soares HD, Chen Y, Sabbagh M, Rohrer A, Schrijvers E, et al. (2009) Identifying Early Markers of Alzheimer's Disease using Quantitative Multiplex Proteomic Immunoassay Panels. *Annals of the New York Academy of Sciences* 1180: 56–67.
- Rocha de Paula M, Gómez Ravetti M, Berretta R, Moscato P (2011) Differences in abundances of cell-signalling proteins in blood reveal novel biomarkers for early detection of clinical Alzheimer's disease. *PLoS ONE* 6(3): e17481. doi:10.1371/journal.pone.0017481.
- Marksteiner J, Kemmler G, Weiss EM, Knaus G, Ullrich C, et al. (2011) Five out of 16 plasma signaling proteins are enhanced in plasma of patients with mild cognitive impairment and Alzheimer's disease. *Neurobiology of aging* 3: 539–40.
- Knudsen LS, Christensen IJ, Lottenburger T, Svendsen MN, Nielsen L, et al. (2008) Pre-analytical and biological variability in circulating interleukin 6 in healthy subjects and patients with rheumatoid arthritis. *Biomarkers* 13: 59–78.
- Ioannidis JP, Panagiotou OA (2011) Comparison of effect sizes associated with biomarkers reported in highly cited individual articles and in subsequent meta-analyses. *JAMA* 21: 2200–10.