AIMS Genetics, 3(1): 25-48.
DOI: 10.3934/genet.2016.1.25
Received 15 January 2016,
Accepted 17 March 2016,
Published 21 March 2016

http://www.aimspress.com/journal/Genetics

Review

Towards defining the Mechanisms of Alzheimer's disease based on a contextual analysis of molecular pathways

Joanna L. Richens[1], Jonathan P. Bramble[1], Hannah L. Spencer[1], Fiona Cantlay[1], Molly Butler[1], and Paul O'Shea[1,2,*]

[1] Cell Biophysics Group, School of Life Sciences, University of Nottingham, University Park, Nottingham, United Kingdom
[2] Address as of 1st July 2016: Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, Canada

* **Correspondence:** Email: paul.oshea@nottingham.ac.uk, poshea@mail.ubc.ca; Tel: +44(0)1159513209; Fax: +44(0)1158466580.

Abstract: Alzheimer's disease (AD) is posing an increasingly profound problem to society. Our genuine understanding of the pathogenesis of AD is inadequate and as a consequence, diagnostic and therapeutic strategies are currently insufficient. The understandable focus of many studies is the identification of molecules with high diagnostic utility however the opportunity to obtain a further understanding of the mechanistic origins of the disease from such putative biomarkers is often overlooked. This study examines the involvement of biomarkers in AD to shed light on potential mechanisms and pathways through which they are implicated in the pathology of this devastating neurodegenerative disorder. The computational tools required to analyse ever-growing datasets in the context of AD are also discussed.

Keywords: dementia; disease pathways; VisANT; biomarkers; clinical bioinformatics

1. Disease-defining biomarkers

Clinical practice throughout history has always been based on the identification and treatment of medical indicators [1]. Rapid advances in scientific knowledge since the 19th century heralded an era of more systematic medical approaches with the development and application of techniques such as anaesthesia [2], aseptic surgery [3] and vaccination [4]. During this time physicians also began to

analyse more systematically and objectively patients' symptoms, relying less on verbal communication and more on the measurable symptoms and causes of disease as demonstrated by the explosion in the field of medical bacteriology [5]. Continual advances in laboratory techniques and equipment over the last century mean that modern day medicine now relies heavily on the identification and quantification of biological markers termed "biomarkers". According to the Biomarkers Definitions Working Group a biomarker can be defined as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention" [6]. The recent rapid expansion in the fields of genomics and proteomics has resulted in many studies referring interchangeably between the term biomarker and molecular entities associated with disease i.e. genetic mutations or aberrantly expressed proteins, but the definition also encompasses measurements such as blood pressure readings for stroke [7], spirometry readings i.e. the Forced Expiratory Volume in 1 second (FEV1)/Forced Vital Capacity (FVC) ratio for Chronic Obstructive Pulmonary Disease (COPD) [8], advanced imaging technologies e.g. single photon emission computed tomography for Parkinson's disease [9] and the urine sediment examination for kidney disease [10].

An idealised biomarker would be one which is only present in either the normal or varied state as this would provide a definitive answer during diagnosis. One such biomarker is the human chorionic gonadotropin (hCG) hormone utilised in pregnancy tests [11]. Such examples are relatively unusual however and many biomarker studies focus on the under- or over-expression of molecules as potential targets. One success story of this approach has been the identification of the membrane tyrosine kinase receptor HER2 as a biomarker for approximately 20–25% of breast cancer cases [12,13]. This finding led to the development of the anti-HER2 humanised monoclonal antibody Trastuzumab (Herceptin™) and has greatly improved the outcome for many patients [13].

2. The difficulties associated with biomarker identification

Over the last two decades numerous putative biomarkers have been identified for many different diseases but gaining the necessary levels of clinical confidence from a single biomarker is not a trivial task. Despite the large financial investment in the field the FDA is approving less than 1.5 new tests per year [14]. One of the biggest limitations to successful biomarker studies is the nature and availability of the clinical sample under investigation. Obtaining a reliable, well characterised collection of samples of known origin can be difficult and limited sample access can sometimes lead to researchers performing studies on samples of unknown integrity [14]. Any sub-optimal treatment that the samples may have been subjected to during collection, processing and storage [15] or any deviation from the expected protocol (e.g. different anti-coagulants or time of day at which the sample was collected) would impact on the quality of the experimental data collected and could result in false-positive biomarker identification. Limited sample access also restricts the size of the cohorts that can be examined. This, along with the cost and time required to process and analyse samples can lead to an under-powering of studies. As a consequence the validation of many biomarkers identified is often inadequate [14].

The optimal clinical sample for ease of routine patient screening and testing would be blood plasma or serum [15]. Plasma and serum are complex fluids comprising a multitude of proteins and metabolites [16,17]. The large dynamic range of these components makes demands on the analytical technologies used during sample analysis and has resulted in the development of methods to try and evade these difficulties [18]. Potentially the biggest limitation to successful biomarker identification

in complex fluids is, however, the huge variation observed in "normal" biological concentrations of proteins and metabolites between individuals [19]. Even when the concentration of a prospective biomarker is shown to be significantly different in the disease condition compared to that in the healthy, there still remains uncertainty as to the accuracy of the clinical decision based on the marker measurement from an individual patient. This uncertainty has its origins in the overlap between the marker distributions (Figure 1). Only those concentration values falling outside of the overlapping ranges can be considered as an affirmative diagnosis, thus extensive levels of overlap may render the practical application of a marker untenable. For those biomarkers that do get adopted into clinical practice the need to establish cut-off points for diagnostic tests can result in the occurrence of false positive and false negative results as exemplified by prostate specific antigen (PSA) a biomarker that is analysed routinely in the diagnostic protocol for prostate cancer but which has come under scrutiny recently [20].

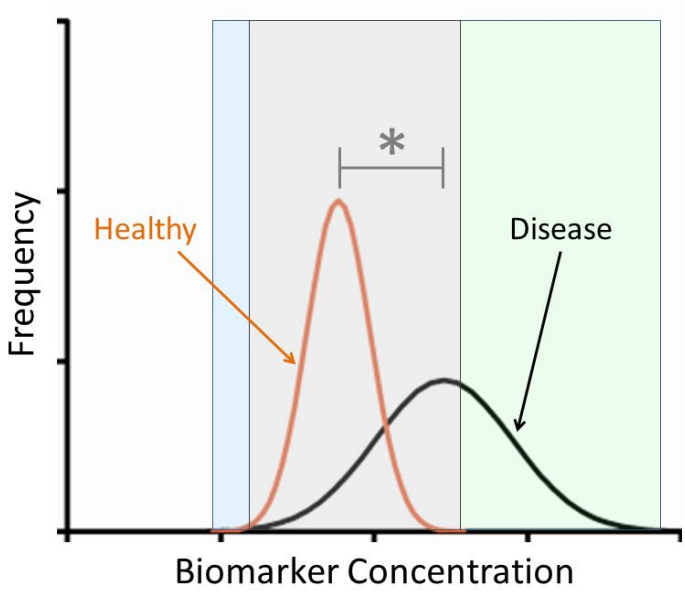

Figure 1. Individual biomarkers in disease detection. Whilst there may be a significant difference (*) between the mean concentrations of biomarker in the healthy and disease populations this does not necessarily deem it appropriate for use in a clinical setting. In this exemplar, the majority of both the disease and healthy populations fall within a grey area whereby the concentration of biomarker measured in their clinical sample could not be used to determine the disease state of the individual.

3. Disease profiling

Determining accurate and reliable biomarkers is critical for the prediction, identification, classification, monitoring and treatment of all diseases. Alongside this, however, biomarkers also offer the often overlooked opportunity to obtain a further understanding of the mechanistic origins of a disease. A biomarker informs us about disease for a reason; significant deviance from a "normal" measurement is either causative of the disease or a consequence of the disease process. There has been a tendency to focus on defining single targets for any one disease. The majority of diseases are, however, highly complex in nature and have multiple factors contributing to their onset and

progression making it unlikely that one biomarker will suffice for all requirements of a biomarker [19]. One way to address this would be to instead examine a variety of parameters that allow a picture of an individual's disease status to be constructed. This patient profiling approach would result in a disease fingerprint comprising genetic, molecular and other physiological, psychological and anatomical components a lot of which would have been identified as putative biomarkers in previous studies. By compiling this information and determining the biological pathways underlying the dysregulation of these targets we would obtain a better understanding of disease pathology. The resultant disease profiles could then also be used to inform future diagnostic and drug development protocols.

4. Alzheimer's disease

Dementia, a syndrome caused by neurodegenerative diseases, is posing a huge societal problem. By 2040 it is predicted that 90.3 million individuals will be affected worldwide [21] with Alzheimer's disease (AD) accounting for approximately 60% of all cases. AD is a complex, progressive degenerate health condition which results in irreversible neuronal damage and a subsequent decrease in cognitive, behavioural and functional abilities [22]. It can be classified into two forms; early-onset AD (EOAD) and late-onset AD (LOAD). LOAD is the most common form of dementia whereas EOAD only accounts for around 1–6% of AD cases [23]. EOAD is defined as AD occurring before the age of 65, and in the case of familial AD (FAD), has a high degree of heritability being strongly associated with mutations in three different genes: amyloid precursor protein (APP) and Presenilins 1 and 2 (PS1/2) [24]. Together, these mutations lead to abnormal processing of APP, resulting in pathogenic Aβ accumulation [24]. Conversely, LOAD typically occurs after the age of 65 and whilst genetics plays a significant role in its onset [25], it is a complex disease and does not possess an autosomal dominant inheritance pattern. This, along with other emerging evidence, suggests that LOAD and FAD are two etiologically different diseases [26]. The multifactorial nature of LOAD means that the underlying disease mechanisms are likely to be extremely complicated and unlikely to be caused solely by amyloid and tau based systems. The remainder of this review focuses, therefore, on the late-onset manifestation of AD.

Existing procedures for the diagnosis of LOAD leave much to be desired and at present it can only be definitively diagnosed post-mortem [25]. Clinical decision-making relies on a complex interplay of clinical and neuropsychometric evaluation and as such it requires significant time investment from experienced clinicians and healthcare professionals with patients before a diagnosis is reached. On a per-patient-basis, this represents an expensive process and could also be considered relatively subjective, relying on the skills, experience and abilities of clinicians forming a diagnostic opinion. As a consequence, existing diagnostic approaches succeed best in moderate to severe stages of dementia. There is however, a much poorer diagnostic sensitivity and selectivity in the early-phase disease whereas the impact of (future) treatment could be most profound at these pre-symptomatic (cognitive) stages [27].

There is a wealth of literature proposing different molecules as putative predictive and prognostic LOAD biomarkers but translational success has been limited. The amyloid cascade hypothesis has been central to AD research since it was first postulated in 1991 [28]. This theory proposes that abnormal formation of Aβ plaques in the brain triggers a sequence of events including tau phosphorylation and neurofibrillary tangle formation that eventually result in synaptic dysfunction, memory loss and structural brain damage [29]. As a consequence, much of the research

effort into LOAD protein biomarkers has focused on the amyloid beta peptides (Aβ), in particular Aβ42, and tau with limited success [30].

APP is an integral membrane protein that comprises a large extracellular domain, a membrane anchoring domain and a short intracellular C-terminal tail [31]. Cleavage of the protein by any one of several α-secretases results in production of soluble APP which is thought to play an important neurotrophic role in facilitating normal physiological functions [32]. Sequential cleavage of APP by firstly β-secretase and then γ-secretase results in production of Aβ peptides which vary in length from 39–43 amino acids [31]. Approximately 90% of the Aβ peptides produced are Aβ40 whilst less than 10% are Aβ42, however, it is the Aβ42 form which is the more toxic peptide due in part to its tendency to aggregate and form fibrils [32]. Soluble Aβ42 oligomers have been implicated in a range of neurotoxic effects including synaptic density reduction, disruption of synaptic transmission, inhibition of hippocampal long-term potentiation and impairment of cognitive functions [32]. However, Aβ42 has also been shown to have some neuroprotective properties at low concentrations when it is unlikely to exist in an oligomeric form [33].

Tau is a microtubule associated protein found predominantly in the axons of neurons which plays an important role in promoting the assembly and stability of microtubules [34]. In its native form it is unfolded, highly soluble and demonstrates very little aggregation, but in tauopathies such as AD aggregation of tau results in formation of neurofibrillary tangles [35]. The normal physiological functioning of tau is highly regulated by a number of post-translational modifications including phosphorylation at some of its 80 serine/threonine and 5 tyrosine phosphorylation sites. In the brains of healthy individuals tau contains approximately 2–3 moles of phosphate per mole of protein, however, in an AD brain tau is hyperphosphorylated to at least three-fold that level [36]. This hyperphosphorylation is believed to be a prerequisite of neurofibrillary tangle formation.

Although amyloid and tau are undeniably linked to LOAD, the nature and significance of this involvement is unclear [37,38]. Whilst mutations within the APP gene are causative of EOFAD [39], there is little evidence for a genetic link between the APP gene and LOAD. There is no reported evidence of a significant correlation between amyloid plaque deposition and either severity of AD or degeneration of neurons [40]. In fact, the amount of plaque formation found in the brain of an AD patient is often similar to that found in a normal individual [41]. It is also unclear as to whether plaques are a cause or a consequence of the disease process and the stage at which they or the preceding Aβ peptides are pathogenic [42,43]. In animal models of AD abnormalities have been detected prior to amyloid plaque formation which may have been due to the toxicity of abnormally high levels of free Aβ peptide [44,45]. Therapeutic approaches focused upon disruption of amyloid-based processes have had limited effect; failing to improve cognition or decrease the rate of mental deterioration and in some cases even seeming to cause enhanced disease progression [46,47]

Regardless of the nature of the contribution of Aβ and tau to LOAD, is becoming apparent that the disease comprises far more than just these two molecules [48] and actually incorporates multiple genetic, molecular and environmental components [49,50]. The multifactorial nature of LOAD along with its close resemblance to other neurological disorders such as vascular dementia, frontotemporal dementia and dementia with Lewy bodies and the need to distinguish it from EOFAD and mild cognitive impairment (MCI) make it a prime candidate for a disease profiling approach. In the next part of this work we explore how such an approach may be applied to LOAD in the context of current diagnostic and clinical knowledge and discuss how re-evaluating how we interpret this knowledge may help to further our understanding of the mechanistic origins of the disease.

5. Constructing a profile of Alzheimer's disease

In order to construct an accurate profile of an individual with dementia it is important to consider a wide variety of parameters which may contribute to the disease phenotype (Figure 2). The most prolific area of recent development in the understanding of LOAD has been at a molecular level. It is by deciphering the interactions and pathways that occur at this level that we will gain a better understanding of the causes of disease onset and begin to comprehend how and why the disease progresses. As such, the basis of a disease profile will comprise predominantly molecular components. These quantitative measurements could be supplemented with information on other disease risk factors including environmental and lifestyle factors, anatomical profiling (e.g. MRI scans) and the psychological profiling that currently underpins a dementia assessment.

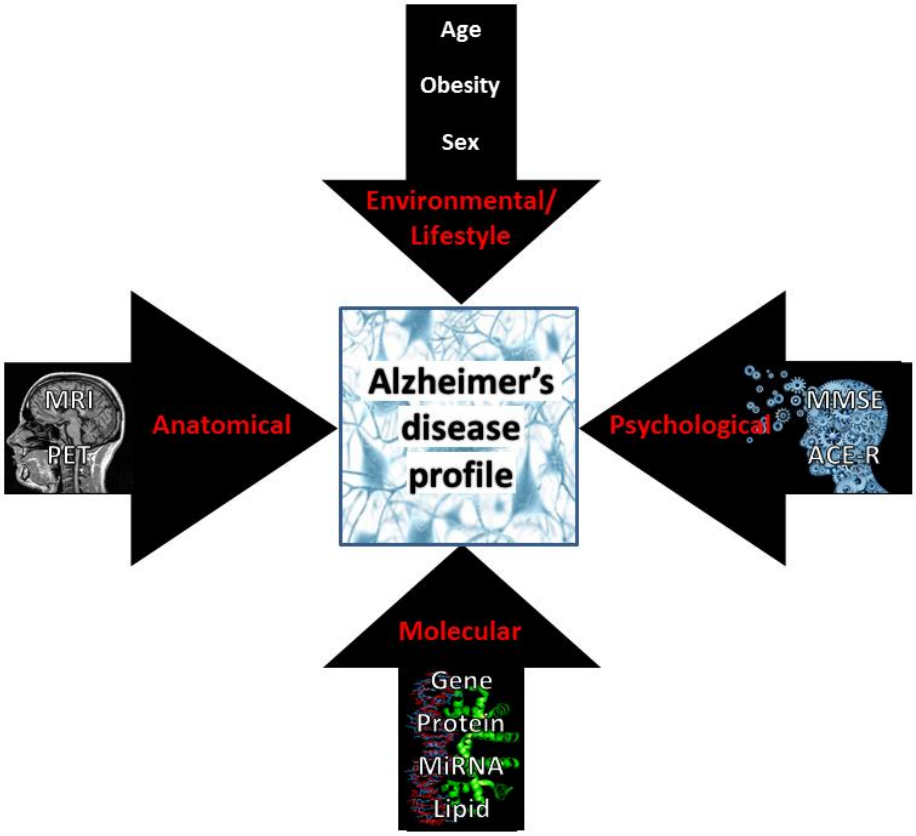

Figure 2. The component parts of an Alzheimer's disease profile.

6. Environmental and lifestyle factors

Whilst environmental and lifestyle factors such as diet, exercise, sleep patterns, smoking, alcohol consumption and education have been shown to influence the onset and progression of LOAD [51], these factors are often qualitative and difficult to enumerate for incorporation into a patient profile. There are, however, several other parameters routinely monitored by physicians which could be considered. Firstly, age is the single greatest risk factor for LOAD [52] with prevalence doubling every five years after the age of 65 [53]. Secondly, gender appears to have some influence on AD development, with evidence suggesting that females are disproportionately affected

by LOAD [54]. Decreased levels of the sex hormones 17β-estradiol and testosterone are observed in the circulation of females and males with LOAD respectively [55] and could be monitored. Whilst testosterone levels decline slowly with age in men, it may be that the sudden decline of estrogens in postmenopausal women contributes to the increased female prevalence of LOAD [54]. There is some evidence linking hypertension with LOAD onset [51] so blood pressure measurements could be included. Finally, mid-life obesity has been identified as a risk factor for LOAD [53] although it seems that central adiposity may be a better predictive indicator to incorporate into a disease profile than body-mass index (BMI) [56]. Interestingly it appears that whilst mid-life obesity can be causative of LOAD an increased BMI may actually be protective in older adults [57].

7. Psychological profiling

Current methods of diagnosing LOAD rely heavily on psychological profiling of the patient by clinical professionals. A selection of cognitive tests for the detection of dementia are available [58] the most regularly used of which is the mini-mental state examination (MMSE) [59]. First published in 1975, this 11-question measure tests five areas of cognitive function: orientation, registration, attention and calculation, recall, and language. The score out of 30 achieved by the patient is indicative of the level of cognitive impairment. An alternative test is the Addenbrooke's cognitive examination-revised (ACE-R) [60] which has been demonstrated to have equal, and in some cases superior, diagnostic accuracy to that of the MMSE [61]. The results from either of these, or any other validated cognitive test could form part of a LOAD profile. There is the possibility however that once accurate quantitative biomarkers have been established for all phases of the disease psychological profiling could be removed from the disease profile to reduce the large time burden currently placed upon clinicians by the AD diagnostic protocol.

8. Anatomical profiling

LOAD is primarily a disease of the brain but biochemical analysis of brain tissue where the characteristic hallmarks, amyloid plaques and neurofibrillary tangles, are located is unfeasible. Advances in imaging technologies mean that it is now possible to gain some insight into the level of brain tissue damage prior to post-mortem examination. A variety of imaging modalities are available including structural and function magnetic resonance imaging (MRI), magnetic resonance spectroscopy (MRS), single photon emission computed tomography (SPECT) and positron emission tomography (PET). A detailed review of the implementation of these technologies in LOAD diagnostics is available [62]. Application of imaging technologies provides important information on temporal dynamic changes of the brain due to LOAD and the spatial distribution of tissue damage.

Studies have demonstrated that implementing imaging technologies alongside cognitive assessment can increase the accuracy of LOAD diagnosis. It is difficult, however, to envisage their widespread, routine adoption in population screening for dementia. Such protocols are time consuming, qualitative and require highly skilled personnel to undertake them. This, along with the expense associated with the manufacture, installation and maintenance of the necessary instrumentation results in high costs per patient scan. Instrumentation is also of limited availability outside specialist centres. Thus, in their currents forms, it is more probable that they will continue to be utilised within a diagnostic programme which includes other schemes e.g. biochemical psychological profiling.

9. Molecular analysis

9.1. Genomic

APOE was the first gene to be identified as a LOAD risk factor [63,64] and for many years it remained the only known genetic indicator of the disease. Three common alleles of the gene exist; ε2, ε3 and ε4. These encode three isoforms of the ApoE protein, a 299 amino residue glycoprotein which vary at positions 112 and 158: ε2 (Cys112, Cys 158), ε3 (Cys 112, Arg158), ε4 (Arg112, Arg158) [65]. Numerous studies have demonstrated an association between the APOE ε4 allele and risk of LOAD and a meta-analysis of some of this work demonstrated an increased risk of nearly fourfold when compared to the ε3 allele [66]. Conversely, carriers of the ApoE ε2 allele appear to have some level of protection against the development of LOAD [67]. As with many disease susceptibility genes, however, *APOE* genotype alone cannot determine disease status. Approximately half of the individuals who develop LOAD are not carriers of the APOE ε4 allele and of those individuals who are carriers, it is estimated that only approximately 10% will develop the disease [68]. Reliance on it as an individual biomarker would, therefore, result in a high incidence of false positive and negative diagnoses.

The advent of improved genomic technologies and approaches such as genome-wide association studies (GWAS) has resulted in an explosion in the number of recognised genetic risk loci to the extent that LOAD is now one of the better genetically characterised diseases. To date at least 25 LOAD genes have been reliably identified [69] and detailed examinations of many of these genetic risk factors are available [70,71]. In addition to *APOE* these genes are *ABCA7* (ATP-binding cassette transporter), *BIN1* (bridging integrator 1), *CASS4* (Cas scaffolding protein family member 4) , *CD2AP* (CD2-associated protein), *CD33* (sialic-acid-binding immunoglobulin-like lectin), *CELF1* locus (CUGBP Elav-like family member 1), *CLU* (clusterin), *CR1* (complement receptor 1), *EPHA1* (ephrin receptor A1), *FERMT2* (Fermitin family homolog 2), *HLA-DRB5*/*HLA-DRB1* locus (Human Leucocyte Antigen DR-beta 1/5), *INPP5D* (Phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase 1), *MEF2C* (Myocyte-specific enhancer factor 2C), *MS4A* cluster (membrane-spanning 4-domains subfamily A), *NME8 (TXNDC3)* locus (Thioredoxin domain-containing protein 3), *PICALM* (phosphatidylinositol-binding clathrin assembly protein), *PTK2B* (Protein tyrosine kinase 2 beta), *SLC24A4*/*RIN3* locus (Sodium/potassium/calcium exchanger 4/Ras and Rab interactor 3), *SORL1* (Sortilin-related receptor), *TREM2* (Triggering receptor expressed on myeloid cells 2) and *ZCWPW1* locus (Zinc finger CW-type PWWP domain protein 1) [72–78].

The identification of so many genetic risk factors of LOAD highlights the multifactorial nature of LOAD. Despite the identification of these new genetic factors of LOAD, however, it is predicted that only approximately 24% of heritability has so far been resolved [10]. As a consequence the implementation of biochemical testing based exclusively on genetic indicators would be unfeasible at the current time as it would result in many false negative diagnoses. The translation of these genetic findings into mechanistic insights into the disease pathology is crucial as biochemical markers may well be independent of genetic markers. Alongside the genetic investigations that have been undertaken, studies into the proteomic aspects of LOAD have also been ongoing. The focus of many of these studies, in some cases unintentionally but at other times as a consequence of the genetic findings, has been on examining the downstream products of the mutated genes. Some of the targets of these genotype-phenotype studies will be discussed in greater detail in the proteomic section of this paper.

9.2. Proteomic

Amyloid plaques and neurofibrillary tau tangles are the pathological hallmarks of AD. As discussed above, whilst there is a large body of evidence implicating both Aβ42 and tau in the pathology of LOAD there is still uncertainty surrounding their application as LOAD biomarkers [62]. Until stronger candidate markers are identified Aβ42 and tau (both phosphorylated and total) are likely to be considered the gold standard and any profiling scheme would include them. The insufficient accuracy of Aβ42 and tau as diagnostic markers for LOAD has, however, led to a widening of investigations into other potential target protein biomarkers.

For many years, mutations in the *APOE* gene were the only known genetic risk factors for LOAD and so the three protein isoforms, ApoE2, ApoE3 and ApoE4 came under much scrutiny. Whilst details of mechanisms underlying the linkage between the APOE ε4 allele and AD still require full clarification, evidence suggests that it contributes to LOAD pathogenesis through both Aβ-dependent and Aβ-independent pathways [79,80]. ApoE co-localises with amyloid plaques [81] and multiple studies have demonstrated ApoE isoform-dependent effects on Aβ clearance and aggregation [80]. In addition, ApoE allele has also been shown to influence neuronal outgrowth and synaptic density and transmission [82,83].

Clusterin (Apolipoprotein J), encoded by the CLU gene on chromosome 8p21-p12, is a chaperone protein widely expressed in brain tissue. A single nucleotide polymorphism (SNP) in the CLU gene has been reproducibly demonstrated to confer a genetic predisposition to LOAD [74,76] but investigations into the involvement of clusterin protein in AD were already ongoing prior to the publication of these findings. Increased expression of clusterin have been shown in the frontal cortex, hippocampus, cerebrospinal fluid (CSF) and plasma of AD patients [84–87] and associations between expression levels and brain atrophy, disease severity and clinical progression have been demonstrated [88,89]. Clusterin binds to both Aβ peptides and amyloid plaques [90,91] and helps to manage levels of soluble and insoluble Aβ in the brain [81] in part through its facilitation of clearance of Aβ across the blood brain barrier as a Clusterin-Aβ complex [92]. Studies on its role in the pathogenesis of AD have indicated that CLU can have both a protective effect by preventing the aggregation of Aβ42, and an exacerbating effect by enhancing the toxicity of Aβ oligomers [93–95].

TREM2 is a glycosylated innate immune receptor, which is thought to be involved in the control of two important signalling cascades in microglial cells: enhancement of phagocytic pathways and suppression of the inflammatory response [73,96]. It has recently gained popularity as a promising biomarker for LOAD as loss of function mutations in this gene have been identified as risk factors for LOAD [73,97]. Additionally, in LOAD patients with the functional TREM2 gene, levels of TREM2 mRNA and protein, and its adaptor protein, DAP12, have been shown to be increased in the brain [98] and its expression pattern appears to map with pathological areas containing amyloid deposits [99,100]. Together, these findings suggest that TREM2 normally has a neuroprotective role against AD, which is possibly triggered by amyloid deposition, resulting in phagocytosis of damaged cells, Aβ clearance and suppression of neuroinflammation. This explains why people with loss of function mutations are more likely to develop LOAD. Interestingly, the signalling pathway initiated by the TREM2/DAP12 complex may also explain several symptoms of LOAD. Although the ligands for this complex are unknown, its activation results in recruitment of spleen tyrosine kinase (syk) [101], which has been implicated in Aβ production and tau hyperphosphorylation [102]. In fact, syk has been shown to directly phosphorylate tau in vitro [103]. Therefore, although an increase in TREM2 levels may initially be neuroprotective, the resulting

increase in syk activation may contribute to the pathological hallmarks of LOAD, suggesting that levels of TREM2 and consequently syk must be tightly controlled as dysregulation either way can have consequences contributing to neurodegeneration in LOAD.

Initial attempts to move proteomic studies away from individual biomarkers towards panels of biomarkers have resulted in tentative successes. Vafadar-Isfahani and colleagues recently identified a panel of CSF biomarkers capable of differentiating healthy individuals from those with AD [104]. This panel comprised Fibrinogen alpha chain precursor (FGA), Keratin type I cytoskeletal 9 (Keratin 9), Serum albumin precursor, SPARC-like 1 protein (SPARCL1) and Tetranectin in addition to the aforementioned Aβ and ApoE. Further studies in our laboratories utilised immunoassays to demonstrate expression of the components of this panel in "matched" CSF and blood plasma samples collected simultaneously from individual donors [105]. The diagnostic power of these markers resided in their combined use as a complete biomarker panel, but the existence of any marker within the panel is suggestive of some involvement in the mechanisms underlying AD pathology.

SPARCL1 is widely expressed in glial and neuronal cells [106]. Little is currently known about the mechanisms through which SPARCL1 is implicated in AD but an association between SPARCL1 and AD has previously been demonstrated by Yin et al. using 1D electrophoresis followed by LC-MS/MS [107]. Evidence also shows a down-regulation of the SPARCL1 gene in hippocampal tissue samples from AD patients when compared to control patients potentially indicating a role for cell adhesion molecules in AD [108]. There is very little other evidence associating tetranectin with AD apart from a study which identified decreased expression levels in the CSF of AD patients [109].

Fibrinogen alpha chain (FGA) is one of three fibrinogen chains, along with beta and gamma fibrinogen which, following cleavage by the protease thrombin, polymerize to form an insoluble fimbrin matrix: one of the major components of blood clots [110]. Several other studies have reported elevated levels of fibrinogens in the brain and blood plasma of AD patients, which accumulate as the disease progresses [111,112]. Although the exact mechanism of its involvement in the disease is unknown, its presence in the AD brain suggests a breakdown of the blood brain barrier as FGA is usually only found in the blood [104,113]. Additionally, it has been shown that FGA specifically binds to Aβ and co-localises with amyloid depositions in the brain [112,114]. The resulting Aβ-fibrin clots are more resistant to degradation [112], which may exacerbate neurodegeneration [113]. Recently, the interaction between Aβ and fibrinogen has been identified as a potential target for AD therapy and initial trials in mice with an inhibitor of the interaction have yielded positive results [113].

Keratin 9 is of particular interest as it was the only component of the biomarker panel that was exclusively detected in the CSF of LOAD patients but not the healthy controls [105]. Furthermore, a previous study by Mueller et al. also used mass spectrometry to demonstrate that Keratin 9 was exclusively present in the blood serum of LOAD patients before cognitive decline but not after [115]. Together these findings suggest that Keratin 9 may be a powerful diagnostic tool for LOAD. This may seem unusual considering Keratin 9 is predominantly expressed in terminally differentiated epidermis of the palms and soles of hands and feet [116]. However, several previous studies [117–122] have identified alternative expression sites of Keratin 9, including serum [122], and a recent study by Li et al. identified Keratin 9 as a component of a protein-protein interaction network in LOAD [123], which substantiates these findings.

9.3. Other molecular targets in blood and CSF

The need to further understand the pathogenic mechanisms and LOAD and the aim of producing a blood-based detection system for the disease has led to investigations into other molecular species and biological processes within blood and CSF. One such species are the micro RNAs (miRNA), 18–25 nt molecules which act as post transcriptional regulators of gene expression by binding to complementary sequences on the 3'UTR of specific RNAs. This results in gene silencing via degradation or translational inhibition [124]. As their level of expression can be influenced by a variety of external factors, they could provide a link between environmental and genetic factors in LOAD onset and progression. MiR-146, which downregulates the inflammatory suppressor protein Complement Factor H (CFH), has been shown to be upregulated in the LOAD brain, contributing to the neuroinflammation that is characteristic of LOAD [125]. Additionally, MiR-195 is thought to be downregulated in the LOAD brain and its normal function is downregulation of the gene encoding BACE1, which reduces amyloidogenic processing of APP and therefore Aβ peptide accumulation [126]. Mir-219 has also been shown to suppress tau synthesis, and its downregulation in the LOAD brain potentially contributes to increased tau toxicity [127].

The lipidome of an individual could also provide some indication of the disease state [65]. Altered levels of lyso-phosphatidylcholine, ceramide and sphingomyelin have all been reported in the CSF of AD patients [128–130] whilst a set of ten targets which were present in the blood plasma of individuals who converted from MCI to LOAD but not in those that did not convert [131] were recently identified. In a separate study, increased levels of monoacylglycerols and diacylglycerols were demonstrated in the frontal cortex and plasma of patients with MCI offering a potential means for earlier disease detection [132].

10. Disease profiling as a tool for understanding disease mechanisms

The molecular targets detailed above are only a sample of the data available in the literature, selected dues to either their significance in the disease process or their diagnostic potential. There are numerous other published markers that have been associated with LOAD but which cannot be detailed individually here. Clinical implementation of many of these targets is often limited due initially to issues with reproducing findings in subsequent studies. The information that they can provide could be invaluable however in furthering our understanding of disease pathogenesis. By connecting these seemingly disparate strands of information together we can begin to collate the details into a unified dataset (disease profile) and subsequently begin to decipher the mechanisms governing their involvement in LOAD. The eventual aim of this approach would be to trace back to the origins of the disease and identify effective therapeutic targets for drug discovery studies.

A range of analytical tools are becoming available that facilitate interrogation of large datasets including Ingenuity (http://www.ingenuity.com), STRING (http://string-db.org) and GeneMANIA (http://www.genemania.org/). These tools enable the biological interactions of an experimentally determined biomarker to be investigated. Examination of multiple targets i.e. a panel of biomarkers, will allow the construction of molecular pathways and networks around these targets. One such software tool which will be used here to exemplify the disease profiling approach is VisANT, an analysis package developed at Boston University and based upon the Predictome database (http://visant.bu.edu) [133]. The majority of interactions utilised by VisANT are experimentally derived and sourced either directly from the literature or from existing database sources such as

BioGRID [134] of the KEGG disease database [133]. In addition to this some interactions are predicted by a robust computational prediction protocol integrated within the Predictome database. VisANT analysis has been applied to a variety of research areas including the mammalian 14-3-3-phosphoproteome [135], the Stat5a network in prostate cancer [136] and the role of FoxP1 in the regulation of autism-related pathways [137].

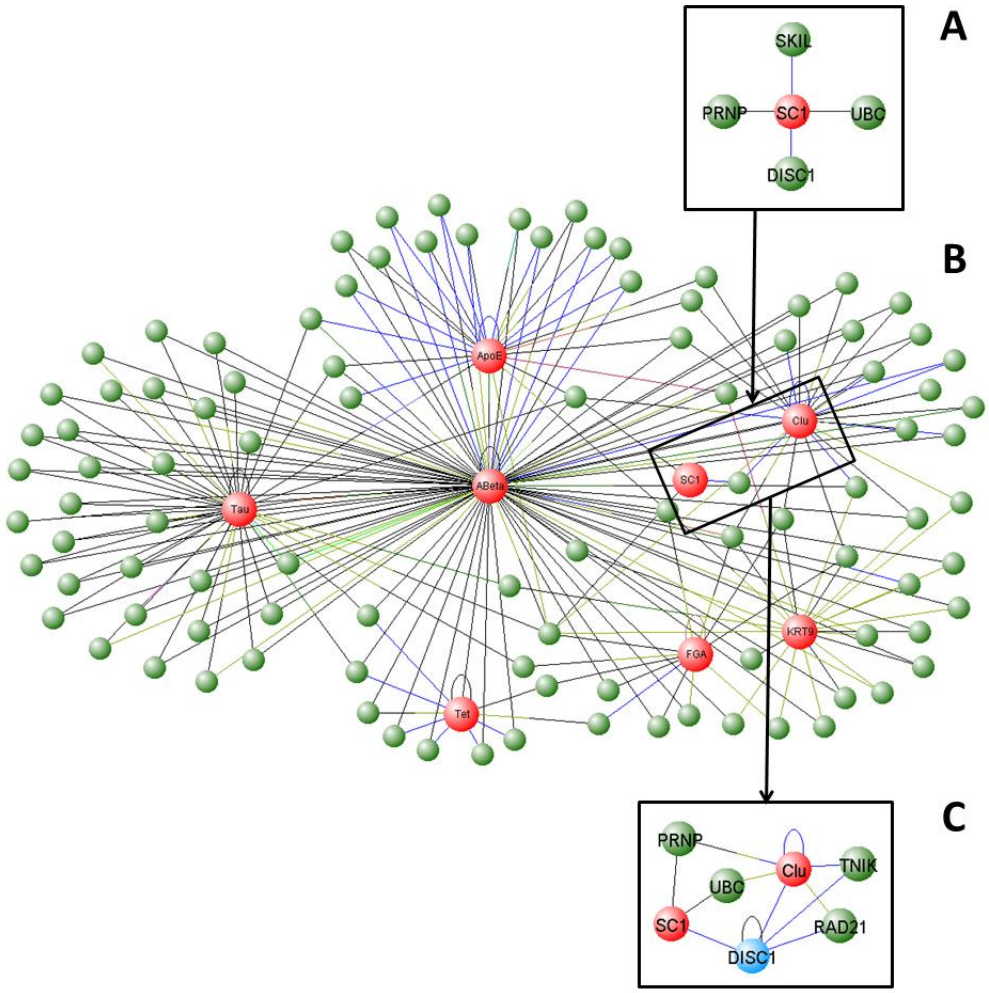

Figure 3. The binding interactions of SPARCL1 (SC1) identified using the VisANT analysis platform (A). When combined with the other putative LOAD protein biomarkers amyloid beta (Aβ), Apolipoprotein E (ApoE), clusterin (Clu), Fibrinogen (FGA), Keratin 9 (KRT9), Tau and Tetranectin (Tet) an interactome is produced (B). Sections of the interest within the interactome can be selected and investigated in more detail as required (C).

In our laboratory we are applying VisANT analysis protocols to AD-related gene and protein panels in order to explore and redefine the mechanistic basis of the disease [49]. The manner in which this problem is being approached can be exemplified using a selection of the proteins detailed above. If we take, for example, SPARCL1 and examine its known interactions it can be seen that it has four reliably determined binding partners; Disrupted in schizophrenia 1 (DISC1), Prion protein (PRNP), SKI-Like Proto-Oncogene (SKIL) and Ubiquitin C (UBC) (Figure 3a). These in themselves

provide some information on the possible linkages between SPARCL1 and LOAD pathogenesis as both DISC1 and PRNP have both been previously implicated in LOAD [138,139]. In order to obtain greater context for the involvement of SPARCL1 in LOAD, it can then be analysed in conjunction with other LOAD factors e.g. Aβ, ApoE, Clusterin, Fibrinogen, Keratin 9, Tau and Tetranectin to produce an interactome (Figure 3b). Figure 3b illustrates a selection of the binding partners and feedback loops which directly connect these molecules. For ease of visualisation, binding partners which only interact with one molecule have been removed from the diagram. The interactome produced from this analysis comprises a total of 108 molecules each of which could be investigated for their contribution to LOAD. Areas of interest can be expanded further to identify greater detail within potential pathways and determine additional molecules of interest (Figure 3c).

Interactome data outputted from VisANT can be further analysed using Reactome (http://www.reactome.org/) [140,141] to define in detail the pathways in which the molecules may be participating. An exemplar of the overview data obtained from Reactome is illustrated in Figure 4. Here the overarching pathway classifications are annotated with the more detailed pathways being diagrammatically illustrated. More detailed examination of each of these pathways can be undertaken as required using Reactome. Connections between different pathways are indicative of molecules which contribute to both pathways. Utilising Reactome allows some level of quantitation to be brought upon the analysis.

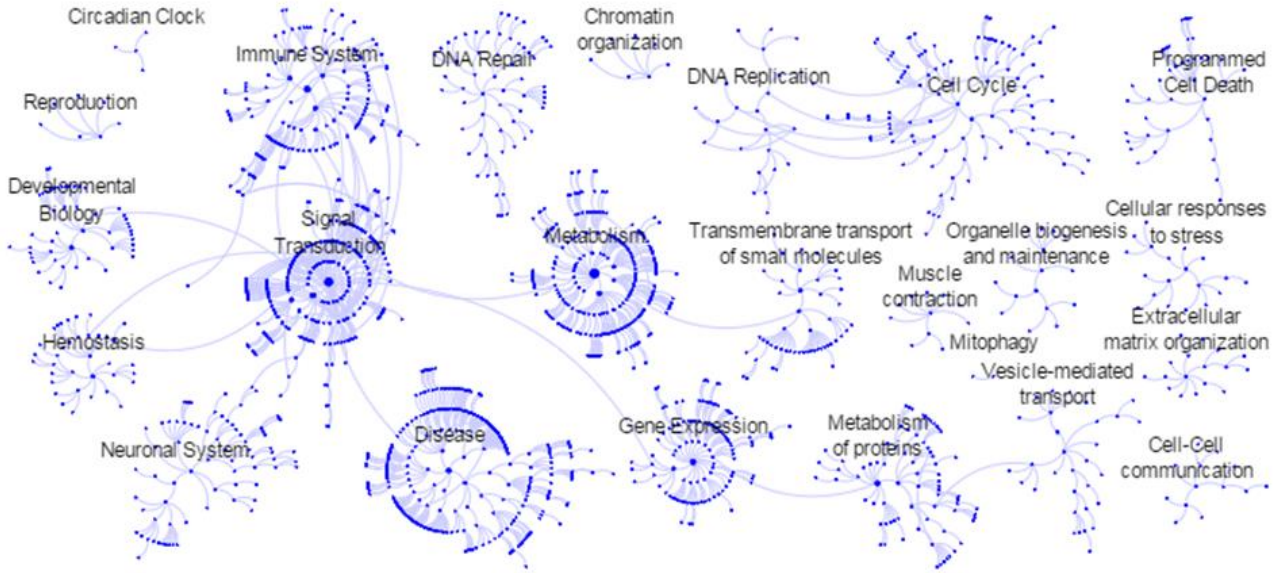

Figure 4. Overview data obtained from Reactome (http://www.reactome.org/) illustrating the broad pathways in which the molecules identified using VisANT participate.

11. Data Science applications and analysis in disease profiling

Big data is a widely used term to describe datasets so large that traditional methods of analysis are unfeasible. This situation has arisen due to the rapid digitisation and collection of data, huge increases in storage capacities and increases in computing power. Associated analysis of these data sets promises to find correlations or trends previously unobserved, to inform business and political decision making. Science is one of the drivers of this technological transformation and medical and

biological fields are prime examples of areas where data collection is both voluminous and varied and where detailed analysis will lead to understanding of disease and positive outcomes for patients.

Clinical Bioinformatics (CBI) combines activities from a number of disciplines such as bioinformatics, medical informatics, data analysis, statistical mathematics and other "omics". It can be defined as the "clinical application of bioinformatics-associated sciences and technologies to understand molecular mechanisms and potential therapies for human diseases" [142]. The aim of CBI is to derive biological and medical information, from multiple sources, to advance the transfer of knowledge towards novel treatments and personal healthcare. The term Translational Bioinformatics has similar definitions and can be considered a synonym for CBI [143,144].

Creation of a personal profile from different data sources requires cooperation between the creators of information to ensure compatible data structures, compatible and secure IT systems. Software and database design would need to be well defined but agile, to enable new data sources to be incorporated in the future. High quality research data from patient studies will be required to provide the basis for disease diagnosis and should be continuously updated to improve reliability.

11.1. Examples of data analysis techniques for diseases diagnosis

Here we outline some recent examples of the applications of CBI to AD and similar conditions, briefly highlighting some of the advanced data analysis methods now employed. Machine Learning is a field of computer science that involves the use of algorithms to study patterns in data, where the computer learns from known data sets and can apply it later to new data. Machine Learning algorithms have become widely used in many areas of computer science and technology including bioinformatics [145]. Here we highlight their use in classification problems for biomarker discovery.

Network analysis has been applied to a wide variety of problems across scientific disciplines in fields such as computer science, chemistry, engineering and bioinformatics. The mathematical basis for these methods, known as graph theory, is well defined and has been implemented in a number of software packages [146]. It has been shown to be useful in the analysis of brain network organisation [147,148], and how structural changes can be determined from experimental measurements. Analysis of this type could form part of a diagnosis method for brain conditions such as AD. In the language of graph theory, vertices represent the brain regions and edges represent the functional connections between regions.

11.2. Data analysis of biomarkers

Screening biomarkers are measured from patient samples using a variety of experimental methods e.g. mass spectrometry, immunoassay and microarray. The aim of the analysis is to identify, from the derived data, which of the markers is a reliable signifier of the disease state. Protein concentrations are reflective of the physiological state of the cells and circulatory fluids within a biological system and can be used to indicate disease. The concentration of these proteins is typically low and we require sensitive techniques to detect and quantify them. Techniques such as mass spectroscopy have relatively low throughput and limit the scope of research studies and their use in regular testing. We require new developments to improve the throughput and sensitivity of these techniques [149].

Pre-processing of raw data to remove outliers, normalise and discretise data and to transform it into the correct formats for further analysis must be performed. Various statistical methods have been employed for outliner removal, based on the assumption of normally distributed data, but other

techniques have been used which may be better suited to biomarker screening [149]. Data pre-processing can determine the success of the analysis and must not be underestimated. Human interaction time in this part of the process can be high [150]. Figure 5 illustrates the processing stages from initial sample collection to biomarker discovery (Figure 5).

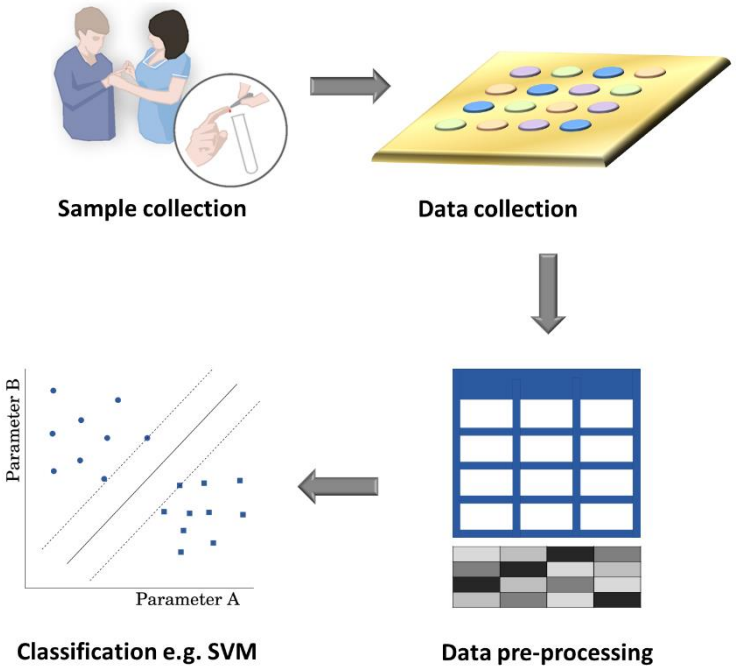

Figure 5. The process of biomarker discovery from sample collection to data classification.

Classified or labelled data from patient studies can be used as training data for machine learning algorithms. Here the algorithm chosen must be able to separate the classifications from sample data, such that a unique region in parameter-space can be identified as an indicator of the disease. It may be that a number of markers must be used simultaneously as an indicator. High-dimensional patterned classifiers for example Support Vector Machines (SVMs) can be used for biomarker selections [145,151]. SVM-like methods have been applied to blood-based biomarkers for the detection of AD [152].

11.3. Data analysis of functional MRI and EEG data

Functional Magnetic Resonance Imaging (fMRI) and positron emission tomography (PET) for neuroimaging have become vital tools in understanding brain structure and function [153]. Recently network analysis has become a successful method of analysing data for the identification of biomarkers of neurodegenerative diseases [153]. Network analysis has been applied to data derived from EEG measurements of patients with AD, Frontotemporal lobar degeneration (FTLD) and no sign of disease [154]. In these experiments parameters derived from readings of functional connectivity, calculated from brain wave patterns, was analysed using graph theory. Various graph parameters such as mean clustering coefficients, characteristic path lengths and degree correlation [146] all were useful parameters in determining disease states. The technique was able to differentiate between patients with AD and FTLD and supported the hypothesis that in AD local brain network

structures breaks down. With a similar analysis approach but with fMRI, Brier et al. [155] studied patients with AD, and determined that the graph parameters of clustering and modularity were reduced in AD patients and potentially could be measured before the onset of symptoms.

11.4. Combining data analysis techniques

Zhang et al. [156] performed combined patient studies with structural MRI, functional PET and quantification of CSF based proteins. They used these complementary techniques as inputs to a multi-kernel SVM for classification. For patients with AD the results when combined together a classification accuracy of 93.2% was achieved, compared to 86.5% for the best individual biomarkers. This demonstrates that machine learning can be used to make classifications from multiple data sources.

12. Conclusion

Putative biomarkers are being continually identified for LOAD but diagnostic and therapeutic advances resulting from these studies have been limited. By utilising the vast amount of knowledge already available to us we can begin to construct biological profiles of the disease. By examining the interactions underlying the profile and the influences conferred by one disease susceptibility factor onto another we can begin to understand mechanisms which govern disease onset and progression. Applications of novel analysis tools such as machine learning will be increasingly necessary to process, classify and extract useful information from these varied and voluminous datasets. Inter-disciplinary collaborations between clinicians, scientists and bioinformaticians must be strengthened to realise the potential benefits of these approaches.

Acknowledgments

This work was funded by a Medical Research Council (MRC) Developmental Pathways Grant (Grant 1001500) and the Engineering and Physical Sciences Research Council (EPSRC) (Grant EP/J017566/1).

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. Strimbu K, Tavel JA (2010) What are Biomarkers? *Curr Opin HIV AIDS* 5: 463-466.
2. Robinson DH, Toledo AH (2012) Historical development of modern anesthesia. *J Invest Surg* 25: 141-149.
3. Pitt D, Aubin JM (2012) Joseph Lister: father of modern surgery. *Can J Surg* 55: E8-9.
4. Hsu JL (2013) A brief history of vaccines: smallpox to the present. *S D Med* Spec no: 33-37.
5. Blevins SM, Bronze MS (2010) Robert Koch and the 'golden age' of bacteriology. *Int J Infect Dis* 14: e744-751.
6. Biomarkers Definitions Working Group (2001) Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 69: 89-95.

7. O'Donnell MJ, Xavier D, Liu L, et al. (2010) Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. *Lancet* 376: 112-123.
8. Richens JL, Urbanowicz RA, Lunt EA, et al. (2009) Systems biology coupled with label-free high-throughput detection as a novel approach for diagnosis of chronic obstructive pulmonary disease. *Respir Res* 10: 29.
9. Song IU, Chung YA, Chung SW, et al. (2014) Early diagnosis of Alzheimer's disease and Parkinson's disease associated with dementia using cerebral perfusion SPECT. *Dement Geriatr Cogn Disord* 37: 276-285.
10. Perazella MA (2015) The Urine Sediment as a Biomarker of Kidney Disease. *Am J Kidney Dis* 66: 748-755.
11. Felton HT, Derrick JB, Swartz DP (1964) A Simple Immunologic Test For Pregnancy. *Can Med Assoc J* 91: 996-1000.
12. Slamon DJ, Clark GM, Wong SG, et al. (1987) Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* 235: 177-182.
13. Puglisi F, Fontanella C, Amoroso V, et al. (2015) Current challenges in HER2-positive breast cancer. *Crit Rev Oncol Hematol* 98: 211-221.
14. Anderson NL, Ptolemy AS, Rifai N (2013) The riddle of protein diagnostics: future bleak or bright? *Clin Chem* 59: 194-197.
15. Mayeux R (2004) Biomarkers: potential uses and limitations. *NeuroRx* 1: 182-188.
16. Anderson NL, Anderson NG (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* 1: 845-867.
17. Krebs HA (1950) Chemical composition of blood plasma and serum. *Annu Rev Biochem* 19: 409-430.
18. Richens JL, Lunt EA, Sanger D, et al. (2009) Avoiding nonspecific interactions in studies of the plasma proteome: practical solutions to prevention of nonspecific interactions for label-free detection of low-abundance plasma proteins. *J Proteome Res* 8: 5103-5110.
19. Drucker E, Krapfenbauer K (2013) Pitfalls and limitations in translation from biomarker discovery to clinical utility in predictive and personalised medicine. *EPMA J* 4: 7.
20. Lim LS, Sherin K (2008) Screening for prostate cancer in U.S. men ACPM position statement on preventive practice. *Am J Prev Med* 34: 164-170.
21. Prince M, Bryce R, Albanese E, et al. (2013) The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimers Dement* 9: 63-75 e62.
22. Galvin JE, Sadowsky CH (2012) Practical guidelines for the recognition and diagnosis of dementia. *J Am Board Fam Med* 25: 367-382.
23. Bekris LM, Yu CE, Bird TD, et al. (2010) Genetics of Alzheimer disease. *J Geriatr Psychiatry Neurol* 23: 213-227.
24. Wu L, Rosa-Neto P, Hsiung GY, et al. (2012) Early-onset familial Alzheimer's disease (EOFAD). *Can J Neurol Sci* 39: 436-445.
25. Ballard C, Gauthier S, Corbett A, et al. (2011) Alzheimer's disease. *Lancet* 377: 1019-1031.
26. Castello MA, Jeppson JD, Soriano S (2014) Moving beyond anti-amyloid therapy for the prevention and treatment of Alzheimer’s disease. *BMC Neurol* 14: 1-5.
27. Cedazo-Minguez A, Winblad B (2010) Biomarkers for Alzheimer's disease and other forms of dementia: clinical needs, limitations and future aspects. *Exp Gerontol* 45: 5-14.

28. Hardy J, Allsop D (1991) Amyloid deposition as the central event in the aetiology of Alzheimer's disease. *Trends Pharmacol Sci* 12: 383-388.
29. Puzzo D, Gulisano W, Arancio O, et al. (2015) The keystone of Alzheimer pathogenesis might be sought in Abeta physiology. *Neuroscience* 307: 26-36.
30. Fiandaca MS, Mapstone ME, Cheema AK, et al. (2014) The critical need for defining preclinical biomarkers in Alzheimer's disease. *Alzheimers Dement* 10: S196-S212.
31. Racchi M, Govoni S (2003) The pharmacology of amyloid precursor protein processing. *Exp Gerontol* 38: 145-157.
32. Nhan HS, Chiang K, Koo EH (2015) The multifaceted nature of amyloid precursor protein and its proteolytic fragments: friends and foes. *Acta Neuropathol* 129: 1-19.
33. Chasseigneaux S, Allinquant B (2012) Functions of Abeta, sAPPalpha and sAPPbeta : similarities and differences. *J Neurochem* 120 Suppl 1: 99-108.
34. Simic G, Babic Leko M, Wray S, et al. (2016) Tau Protein Hyperphosphorylation and Aggregation in Alzheimer's Disease and Other Tauopathies, and Possible Neuroprotective Strategies. *Biomolecules* 6.
35. Wang Y, Mandelkow E (2016) Tau in physiology and pathology. *Nat Rev Neurosci* 17: 22-35.
36. Wang JZ, Xia YY, Grundke-Iqbal I, et al. (2013) Abnormal hyperphosphorylation of tau: sites, regulation, and molecular mechanism of neurofibrillary degeneration. *J Alzheimers Dis* 33 Suppl 1: S123-139.
37. Robakis NK (2010) Are Abeta and its derivatives causative agents or innocent bystanders in AD? *Neurodegener Dis* 7: 32-37.
38. Castello MA, Soriano S (2014) On the origin of Alzheimer's disease. Trials and tribulations of the amyloid hypothesis. *Ageing Res Rev* 13: 10-12.
39. Loy CT, Schofield PR, Turner AM, et al. (2014) Genetics of dementia. *Lancet* 383: 828-840.
40. Arriagada PV, Growdon JH, Hedley-Whyte ET, et al. (1992) Neurofibrillary tangles but not senile plaques parallel duration and severity of Alzheimer's disease. *Neurology* 42: 631-639.
41. Crystal H, Dickson D, Fuld P, et al. (1988) Clinico-pathologic studies in dementia: nondemented subjects with pathologically confirmed Alzheimer's disease. *Neurology* 38: 1682-1687.
42. Benilova I, Karran E, De Strooper B (2012) The toxic A[beta] oligomer and Alzheimer's disease: an emperor in need of clothes. *Nat Neurosci* 15: 349-357.
43. Hardy J (2009) The amyloid hypothesis for Alzheimer's disease: a critical reappraisal. *J Neurochem* 110: 1129-1134.
44. Hsia AY, Masliah E, McConlogue L, et al. (1999) Plaque-independent disruption of neural circuits in Alzheimer's disease mouse models. *Proc Natl Acad Sci U S A* 96: 3228-3233.
45. Mucke L, Masliah E, Yu GQ, et al. (2000) High-level neuronal expression of abeta 1-42 in wild-type human amyloid protein precursor transgenic mice: synaptotoxicity without plaque formation. *J Neurosci* 20: 4050-4058.
46. Holmes C, Boche D, Wilkinson D, et al. (2008) Long-term effects of Abeta42 immunisation in Alzheimer's disease: follow-up of a randomised, placebo-controlled phase I trial. *Lancet* 372: 216-223.
47. Karran E, Mercken M, De Strooper B (2011) The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. *Nat Rev Drug Discov* 10: 698-712.
48. Robakis NK (2011) Mechanisms of AD neurodegeneration may be independent of Abeta and its derivatives. *Neurobiol Aging* 32: 372-379.

49. Richens JL, Morgan K, O'Shea P (2014) Reverse engineering of Alzheimer's disease based on biomarker pathways analysis. *Neurobiol Aging* 35: 2029-2038.
50. Morgan K (2011) The three new pathways leading to Alzheimer's disease. *Neuropathol Appl Neurobiol* 37: 353-357.
51. Baumgart M, Snyder HM, Carrillo MC, et al. (2015) Summary of the evidence on modifiable risk factors for cognitive decline and dementia: A population-based perspective. *Alzheimers Dement* 11: 718-726.
52. Hebert LE, Scherr PA, Bienias JL, et al. (2003) Alzheimer disease in the US population: prevalence estimates using the 2000 census. *Arch Neurol* 60: 1119-1122.
53. Moser VA, Pike CJ (2015) Obesity and sex interact in the regulation of Alzheimer's disease. *Neurosci Biobehav Rev* in press.
54. Li R, Singh M (2014) Sex differences in cognitive impairment and Alzheimer's disease. *Front Neuroendocrinol* 35: 385-403.
55. Barron AM, Pike CJ (2012) Sex hormones, aging, and Alzheimer's disease. *Front Biosci (Elite Ed)* 4: 976-997.
56. Luchsinger JA, Cheng D, Tang MX, et al. (2012) Central obesity in the elderly is related to late-onset Alzheimer disease. *Alzheimer Dis Assoc Disord* 26: 101-105.
57. Emmerzaal TL, Kiliaan AJ, Gustafson DR (2015) 2003-2013: a decade of body mass index, Alzheimer's disease, and dementia. *J Alzheimers Dis* 43: 739-755.
58. Tsoi KK, Chan JY, Hirai HW, et al. (2015) Cognitive Tests to Detect Dementia: A Systematic Review and Meta-analysis. *JAMA Intern Med* 175: 1450-1458.
59. Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 12: 189-198.
60. Mioshi E, Dawson K, Mitchell J, et al. (2006) The Addenbrooke's Cognitive Examination Revised (ACE-R): a brief cognitive test battery for dementia screening. *Int J Geriatr Psychiatry* 21: 1078-1085.
61. Larner AJ, Mitchell AJ (2014) A meta-analysis of the accuracy of the Addenbrooke's Cognitive Examination (ACE) and the Addenbrooke's Cognitive Examination-Revised (ACE-R) in the detection of dementia. *Int Psychogeriatr* 26: 555-563.
62. Hampel H, Frank R, Broich K, et al. (2010) Biomarkers for Alzheimer's disease: academic, industry and regulatory perspectives. *Nat Rev Drug Discov* 9: 560-574.
63. Corder EH, Saunders AM, Strittmatter WJ, et al. (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 261: 921-923.
64. Saunders AM, Strittmatter WJ, Schmechel D, et al. (1993) Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology* 43: 1467-1472.
65. Xiang Y, Lam SM, Shui G (2015) What can lipidomics tell us about the pathogenesis of Alzheimer disease? *Biol Chem* 396: 1281-1291.
66. Bertram L, McQueen MB, Mullin K, et al. (2007) Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet* 39: 17-23.
67. Corder EH, Saunders AM, Risch NJ, et al. (1994) Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nat Genet* 7: 180-184.
68. Myers RH, Schaefer EJ, Wilson PW, et al. (1996) Apolipoprotein E epsilon4 association with dementia in a population-based study: The Framingham study. *Neurology* 46: 673-677.

69. Barber RC, Phillips NR, Tilson JL, et al. (2015) Can Genetic Analysis of Putative Blood Alzheimer's Disease Biomarkers Lead to Identification of Susceptibility Loci? *PLoS One* 10: e0142360.
70. Van Cauwenberghe C, Van Broeckhoven C, Sleegers K (2015) The genetic landscape of Alzheimer disease: clinical implications and perspectives. *Genet Med* in press.
71. Karch CM, Goate AM (2015) Alzheimer's disease risk genes and mechanisms of disease pathogenesis. *Biol Psychiatry* 77: 43-51.
72. Bertram L, Lill CM, Tanzi RE (2010) The genetics of Alzheimer disease: back to the future. *Neuron* 68: 270-281.
73. Guerreiro R, Wojtas A, Bras J, et al. (2013) TREM2 variants in Alzheimer's disease. *N Engl J Med* 368: 117-127.
74. Harold D, Abraham R, Hollingworth P, et al. (2009) Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet* 41: 1088-1093.
75. Hollingworth P, Harold D, Sims R, et al. (2011) Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat Genet* 43: 429-435.
76. Lambert JC, Heath S, Even G, et al. (2009) Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat Genet* 41: 1094-1099.
77. Naj AC, Jun G, Beecham GW, et al. (2011) Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat Genet* 43: 436-441.
78. Seshadri S, Fitzpatrick AL, Ikram MA, et al. (2010) Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA* 303: 1832-1840.
79. Huang Y (2010) Abeta-independent roles of apolipoprotein E4 in the pathogenesis of Alzheimer's disease. *Trends Mol Med* 16: 287-294.
80. Kim J, Yoon H, Basak J, et al. (2014) Apolipoprotein E in synaptic plasticity and Alzheimer's disease: potential cellular and molecular mechanisms. *Mol Cells* 37: 767-776.
81. DeMattos RB, Cirrito JR, Parsadanian M, et al. (2004) ApoE and clusterin cooperatively suppress Abeta levels and deposition: evidence that ApoE regulates extracellular Abeta metabolism in vivo. *Neuron* 41: 193-202.
82. Nathan BP, Bellosta S, Sanan DA, et al. (1994) Differential effects of apolipoproteins E3 and E4 on neuronal growth in vitro. *Science* 264: 850-852.
83. Ji Y, Gong Y, Gan W, et al. (2003) Apolipoprotein E isoform-specific regulation of dendritic spine morphology in apolipoprotein E transgenic mice and Alzheimer's disease patients. *Neuroscience* 122: 305-315.
84. Lidstrom AM, Bogdanovic N, Hesse C, et al. (1998) Clusterin (apolipoprotein J) protein levels are increased in hippocampus and in frontal cortex in Alzheimer's disease. *Exp Neurol* 154: 511-521.
85. Sihlbom C, Davidsson P, Sjogren M, et al. (2008) Structural and quantitative comparison of cerebrospinal fluid glycoproteins in Alzheimer's disease patients and healthy individuals. *Neurochem Res* 33: 1332-1340.
86. Nilselid AM, Davidsson P, Nagga K, et al. (2006) Clusterin in cerebrospinal fluid: analysis of carbohydrates and quantification of native and glycosylated forms. *Neurochem Int* 48: 718-728.
87. Jongbloed W, van Dijk KD, Mulder SD, et al. (2015) Clusterin Levels in Plasma Predict Cognitive Decline and Progression to Alzheimer's Disease. *J Alzheimers Dis* 46: 1103-1110.

88. Thambisetty M, Simmons A, Velayudhan L, et al. (2010) Association of plasma clusterin concentration with severity, pathology, and progression in Alzheimer disease. *Arch Gen Psychiatry* 67: 739-748.
89. Schrijvers EM, Koudstaal PJ, Hofman A, et al. (2011) Plasma clusterin and the risk of Alzheimer disease. *JAMA* 305: 1322-1326.
90. McGeer PL, Kawamata T, Walker DG (1992) Distribution of clusterin in Alzheimer brain tissue. *Brain Res* 579: 337-341.
91. Ghiso J, Matsubara E, Koudinov A, et al. (1993) The cerebrospinal-fluid soluble form of Alzheimer's amyloid beta is complexed to SP-40,40 (apolipoprotein J), an inhibitor of the complement membrane-attack complex. *Biochem J* 293 (Pt 1): 27-30.
92. Zlokovic BV (1996) Cerebrovascular transport of Alzheimer's amyloid beta and apolipoproteins J and E: possible anti-amyloidogenic role of the blood-brain barrier. *Life Sci* 59: 1483-1497.
93. Giannakopoulos P, Kovari E, French LE, et al. (1998) Possible neuroprotective role of clusterin in Alzheimer's disease: a quantitative immunocytochemical study. *Acta Neuropathol* 95: 387-394.
94. Oda T, Wals P, Osterburg HH, et al. (1995) Clusterin (apoJ) alters the aggregation of amyloid beta-peptide (A beta 1-42) and forms slowly sedimenting A beta complexes that cause oxidative stress. *Exp Neurol* 136: 22-31.
95. Li X, Ma Y, Wei X, et al. (2014) Clusterin in Alzheimer's disease: a player in the biological behavior of amyloid-beta. *Neurosci Bull* 30: 162-168.
96. Jiang T, Tan L, Zhu XC, et al. (2014) Upregulation of TREM2 ameliorates neuropathology and rescues spatial cognitive impairment in a transgenic mouse model of Alzheimer's disease. *Neuropsychopharmacology* 39: 2949-2962.
97. Zhao Y, Bhattacharjee S, Jones BM, et al. (2013) Regulation of TREM2 expression by an NF-small ka, CyrillicB-sensitive miRNA-34a. *Neuroreport* 24: 318-323.
98. Lue LF, Schmitz C, Walker DG (2015) What happens to microglial TREM2 in Alzheimer's disease: Immunoregulatory turned into immunopathogenic? *Neuroscience* 302: 138-150.
99. Jiang T, Yu JT, Zhu XC, et al. (2013) TREM2 in Alzheimer's disease. *Mol Neurobiol* 48: 180-185.
100. Frank S, Burbach GJ, Bonin M, et al. (2008) TREM2 is upregulated in amyloid plaque-associated microglia in aged APP23 transgenic mice. *Glia* 56: 1438-1447.
101. Paradowska-Gorycka A, Jurkowska M (2013) Structure, expression pattern and biological activity of molecular complex TREM-2/DAP12. *Hum Immunol* 74: 730-737.
102. Paris D, Ait-Ghezala G, Bachmeier C, et al. (2014) The spleen tyrosine kinase (Syk) regulates Alzheimer amyloid-beta production and Tau hyperphosphorylation. *J Biol Chem* 289: 33927-33944.
103. Lebouvier T, Scales TM, Hanger DP, et al. (2008) The microtubule-associated protein tau is phosphorylated by Syk. *Biochim Biophys Acta* 1783: 188-192.
104. Vafadar-Isfahani B, Ball G, Coveney C, et al. (2012) Identification of SPARC-like 1 Protein as Part of a Biomarker Panel for Alzheimer's Disease in Cerebrospinal Fluid. *J Alzheimers Dis* 28: 625-636.
105. Richens JL, Vere KA, Light RA, et al. (2014) Practical detection of a definitive biomarker panel for Alzheimer's disease; comparisons between matched plasma and cerebrospinal fluid. *Int J Mol Epidemiol Genet* 5: 53-70.

106. Mendis DB, Ivy GO, Brown IR (2000) Induction of SC1 mRNA encoding a brain extracellular matrix glycoprotein related to SPARC following lesioning of the adult rat forebrain. *Neurochem Res* 25: 1637-1644.
107. Yin GN, Lee HW, Cho JY, et al. (2009) Neuronal pentraxin receptor in cerebrospinal fluid as a potential biomarker for neurodegenerative diseases. *Brain Res* 1265: 158-170.
108. Xu PT, Li YJ, Qin XJ, et al. (2006) Differences in apolipoprotein E3/3 and E4/4 allele-specific gene expression in hippocampus in Alzheimer disease. *Neurobiol Dis* 21: 256-275.
109. Abdi F, Quinn JF, Jankovic J, et al. (2006) Detection of biomarkers with a multiplex quantitative proteomic platform in cerebrospinal fluid of patients with neurodegenerative disorders. *J Alzheimers Dis* 9: 293-348.
110. Weisel JW (2005) Fibrinogen and fibrin. *Adv Protein Chem* 70: 247-299.
111. Cortes-Canteli M, Strickland S (2009) Fibrinogen, a possible key player in Alzheimer's disease. *J Thromb Haemost* 7 Suppl 1: 146-150.
112. Cortes-Canteli M, Paul J, Norris EH, et al. (2010) Fibrinogen and beta-amyloid association alters thrombosis and fibrinolysis: a possible contributing factor to Alzheimer's disease. *Neuron* 66: 695-709.
113. Ahn HJ, Glickman JF, Poon KL, et al. (2014) A novel Abeta-fibrinogen interaction inhibitor rescues altered thrombosis and cognitive decline in Alzheimer's disease mice. *J Exp Med* 211: 1049-1062.
114. Ahn HJ, Zamolodchikov D, Cortes-Canteli M, et al. (2010) Alzheimer's disease peptide beta-amyloid interacts with fibrinogen and induces its oligomerization. *Proc Natl Acad Sci U S A* 107: 21812-21817.
115. Mueller C, Zhou W, Vanmeter A, et al. (2010) The heme degradation pathway is a promising serum biomarker source for the early detection of Alzheimer's disease. *J Alzheimers Dis* 19: 1081-1091.
116. Langbein L, Heid HW, Moll I, et al. (1993) Molecular characterization of the body site-specific human epidermal cytokeratin 9: cDNA cloning, amino acid sequence, and tissue specificity of gene expression. *Differentiation* 55: 57-71.
117. Ghosh D, Lippert D, Krokhin O, et al. (2010) Defining the membrane proteome of NK cells. *J Mass Spectrom* 45: 1-25.
118. van Niel G, Raposo G, Candalh C, et al. (2001) Intestinal epithelial cells secrete exosome-like vesicles. *Gastroenterology* 121: 337-349.
119. de Mateo S, Castillo J, Estanyol JM, et al. (2011) Proteomic characterization of the human sperm nucleus. *Proteomics* 11: 2714-2726.
120. Bohm D, Keller K, Pieter J, et al. (2012) Comparison of tear protein levels in breast cancer patients and healthy controls using a de novo proteomic approach. *Oncol Rep* 28: 429-438.
121. Kim YS, Gu BH, Choi BC, et al. (2013) Apolipoprotein A-IV as a novel gene associated with polycystic ovary syndrome. *Int J Mol Med* 31: 707-716.
122. Fu BS, Liu W, Zhang JW, et al. (2009) [Serum proteomic analysis on metastasis-associated proteins of hepatocellular carcinoma]. *Nan Fang Yi Ke Da Xue Xue Bao* 29: 1775-1778.
123. Li X, Long J, He T, et al. (2015) Integrated genomic approaches identify major pathways and upstream regulators in late onset Alzheimer's disease. *Sci Rep* 5: 12393.
124. Femminella GD, Ferrara N, Rengo G (2015) The emerging role of microRNAs in Alzheimer's disease. *Front Physiol* 6: 40.

125. Lukiw WJ, Zhao Y, Cui JG (2008) An NF-kappaB-sensitive micro RNA-146a-mediated inflammatory circuit in Alzheimer disease and in stressed human brain cells. *J Biol Chem* 283: 31315-31322.
126. Zhu HC, Wang LM, Wang M, et al. (2012) MicroRNA-195 downregulates Alzheimer's disease amyloid-beta production by targeting BACE1. *Brain Res Bull* 88: 596-601.
127. Santa-Maria I, Alaniz ME, Renwick N, et al. (2015) Dysregulation of microRNA-219 promotes neurodegeneration through post-transcriptional regulation of tau. *J Clin Invest* 125: 681-686.
128. Kosicek M, Zetterberg H, Andreasen N, et al. (2012) Elevated cerebrospinal fluid sphingomyelin levels in prodromal Alzheimer's disease. *Neurosci Lett* 516: 302-305.
129. Mulder C, Wahlund LO, Teerlink T, et al. (2003) Decreased lysophosphatidylcholine/phosphatidylcholine ratio in cerebrospinal fluid in Alzheimer's disease. *J Neural Transm (Vienna)* 110: 949-955.
130. Satoi H, Tomimoto H, Ohtani R, et al. (2005) Astroglial expression of ceramide in Alzheimer's disease brains: a role during neuronal apoptosis. *Neuroscience* 130: 657-666.
131. Mapstone M, Cheema AK, Fiandaca MS, et al. (2014) Plasma phospholipids identify antecedent memory impairment in older adults. *Nat Med* 20: 415-418.
132. Wood PL, Medicherla S, Sheikh N, et al. (2015) Targeted Lipidomics of Fontal Cortex and Plasma Diacylglycerols (DAG) in Mild Cognitive Impairment and Alzheimer's Disease: Validation of DAG Accumulation Early in the Pathophysiology of Alzheimer's Disease. *J Alzheimers Dis* 48: 537-546.
133. Hu Z, Chang YC, Wang Y, et al. (2013) VisANT 4.0: Integrative network platform to connect genes, drugs, diseases and therapies. *Nucleic Acids Res* 41: W225-231.
134. Chatr-Aryamontri A, Breitkreutz BJ, Oughtred R, et al. (2015) The BioGRID interaction database: 2015 update. *Nucleic Acids Res* 43: D470-478.
135. Johnson C, Tinti M, Wood NT, et al. (2011) Visualization and biochemical analyses of the emerging mammalian 14-3-3-phosphoproteome. *Mol Cell Proteomics* 10: M110 005751.
136. Reddy A, Huang CC, Liu H, et al. (2010) Robust gene network analysis reveals alteration of the STAT5a network as a hallmark of prostate cancer. *Genome Inform* 24: 139-153.
137. Araujo DJ, Anderson AG, Berto S, et al. (2015) FoxP1 orchestration of ASD-relevant signaling pathways in the striatum. *Genes Dev* 29: 2081-2096.
138. Haas LT, Salazar SV, Kostylev MA, et al. (2015) Metabotropic glutamate receptor 5 couples cellular prion protein to intracellular signalling in Alzheimer's disease. *Brain* 139: 526-546.
139. Shahani N, Seshadri S, Jaaro-Peled H, et al. (2015) DISC1 regulates trafficking and processing of APP and Abeta generation. *Mol Psychiatry* 20: 874-879.
140. Croft D, O'Kelly G, Wu G, et al. (2011) Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res* 39: D691-697.
141. D'Eustachio P (2013) Pathway Databases: Making Chemical and Biological Sense of the Genomic Data Flood. *Chem Biol* 20: 629-635.
142. Wang X, Liotta L (2011) Clinical bioinformatics: a new emerging science. *J Clin Bioinforma* 1: 1.
143. Bellazzi R, Masseroli M, Murphy S, et al. (2012) Clinical Bioinformatics: challenges and opportunities. *BMC Bioinform* 13: 1-8.
144. Butte AJ (2008) Translational Bioinformatics: Coming of Age. *J Am Med Inf Assoc* 15: 709-714.
145. Larrañaga P, Calvo B, Santana R, et al. (2006) Machine learning in bioinformatics. *Brief Bioinform* 7: 86-112.

146. Pavlopoulos GA, Secrier M, Moschopoulos CN, et al. (2011) Using graph theory to analyze biological networks. *BioData Min* 4: 1-27.
147. Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* 10: 186-198.
148. Stam CJ (2014) Modern network science of neurological disorders. *Nat Rev Neurosci* 15: 683-695.
149. Baumgartner C, Osl M, Netzer M, et al. (2011) Bioinformatic-driven search for metabolic biomarkers in disease. *J Clin Bioinforma* 1: 1-10.
150. Inza I, Calvo B, Armañanzas R, et al. (2010) Machine Learning: An Indispensable Tool in Bioinformatics. In: Matthiesen R, editor. *Bioinformatics Methods in Clinical Research*: Humana Press. pp. 25-48.
151. Guyon I, Weston J, Barnhill S, et al. Gene Selection for Cancer Classification using Support Vector Machines. *Mach Learn* 46: 389-422.
152. Hsu W-C, Denq C, Chen S-S (2013) A diagnostic methodology for Alzheimer's disease. *J Clin Bioinforma* 3: 1-13.
153. Horwitz B, Rowe JB (2011) Functional biomarkers for neurodegenerative disorders based on the network paradigm. *Prog Neurobiol* 95: 505-509.
154. de Haan W, Pijnenburg YAL, Strijers RLM, et al. (2009) Functional neural network analysis in frontotemporal dementia and Alzheimer's disease using EEG and graph theory. *BMC Neurosci* 10: 101-101.
155. Brier MR, Thomas JB, Fagan AM, et al. (2014) Functional connectivity and graph theory in preclinical Alzheimer's disease. *Neurobiol Aging* 35: 757-768.
156. Zhang D, Wang Y, Zhou L, et al. (2011) Multimodal classification of Alzheimer's disease and mild cognitive impairment. *NeuroImage* 55: 856-867.

AIMS Press