Raman spectroscopy to diagnose Alzheimer's disease and dementia with Lewy bodies in blood

Maria Paraskevaidi[a,*], Camilo L. M. Morais[a], Diane E. Halliwell[a], David M. A. Mann[b], David Allsop[c], Pierre L. Martin-Hirsch[d] and Francis L. Martin[a,*]

[a]School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston PR1 2HE, UK

[b]Division of Neuroscience and Experimental Psychology, School of Biological Sciences, University of Manchester, Greater Manchester Neurosciences Centre, Salford Royal Hospital, Salford M6 8HD, UK

[c]Division of Biomedical and Life Sciences, Faculty of Health and Medicine, Lancaster University, Lancaster LA1 4YQ, UK

[d]Department of Obstetrics and Gynaecology, Central Lancashire Teaching Hospitals NHS Foundation Trust, Preston PR2 9HT, UK

[*]To whom correspondence should be addressed: mparaskevaidi@uclan.ac.uk or flmartin@uclan.ac.uk

Abstract

Accurate identification of Alzheimer's disease (AD) is still of major clinical importance considering the current lack of non-invasive and low-cost diagnostic approaches. Detection of early-stage AD is particularly desirable as it would allow early intervention and/or recruitment of patients into clinical trials. There is also an unmet need for discrimination of AD from dementia with Lewy bodies (DLB), as many cases of the latter are misdiagnosed as AD. Biomarkers based on a simple blood test would be useful in research and clinical practice. Raman spectroscopy has been implemented to analyse blood plasma of a cohort that consisted of early-stage AD, late-stage AD, DLB and healthy controls. Classification algorithms achieved high accuracy for the different groups: early-stage AD *vs* healthy with 84% sensitivity, 86% specificity; late-stage AD *vs* healthy with 84% sensitivity, 77% specificity; DLB *vs* healthy with 83% sensitivity, 87% specificity; early-stage AD *vs* DLB with 81% sensitivity, 88% specificity; late-stage AD *vs* DLB with 90% sensitivity, 93% specificity; and lastly, early-stage AD *vs* late-stage AD 66% sensitivity and 83% specificity. G-score values were also estimated between 74-91%, demonstrating that the overall performance of the classification model was satisfactory. The wavenumbers responsible for differentiation were assigned to important biomolecules which can serve as a panel of biomarkers. These results suggest a cost-effective, blood-based biomarker for neurodegeneration in dementias.

Keywords: Alzheimer's disease; Dementia with Lewy bodies; Raman spectroscopy; blood plasma; biomarkers

Introduction

Alzheimer's disease (AD) and dementia with Lewy bodies (DLB) constitute the two most common causes of dementia. AD and DLB can share common symptoms and clinical characteristics, which can lead to misdiagnosis. A clear distinction between these two causes of dementia is necessary in terms of pharmacological treatment and outcome evaluation [1, 2]. The neuropathological hallmarks of AD include senile plaques (containing accumulated amyloid-beta (Aβ) peptide) and neurofibrillary tangles (composed of hyperphosphorylated tau protein), while in DLB the hallmark pathology is the abnormal aggregation of α-synuclein into Lewy bodies and Lewy neurites [3, 4]. The ability to index the presence of these pathological features in very early stages (*i.e.*, prodromal disease), or even before symptoms occur (*i.e.*, pre-clinical disease), would allow an earlier intervention before irreversible neuronal death occurs, as well as facilitating early recruitment into clinical trials.

Accurate detection of dementia is essential for improving the lives of those affected. Current diagnostic approaches employ neuroimaging techniques, such as magnetic resonance imaging (MRI) and positron emission tomography (PET) scans (amyloid-PET and more recently tau-PET), or cerebrospinal fluid (CSF) biomarkers, but these methods have many limitations [5-8]. A combination of family and clinical history, as well as a series of different memory and psychological tests is often required for diagnosis, but not all pathologically similar cases will present with the same "clinical phenotype"; many studies have shown contradictory results regarding the suitability of these biomarkers for accurate diagnosis. Recently, blood biomarkers have emerged as a potential means to test for neurodegenerative diseases, with some being capable of detecting early-stage disease [9-11]. The rationale behind the use of blood samples is based on the daily release of 500 ml CSF into the bloodstream, which potentially renders blood a rich source of brain biomarkers [12].

Raman is a spectroscopic technique that extracts biological information by applying a monochromatic, laser light onto the sample under interrogation; electrons are thus excited to virtual energy levels. When these electrons return to the original energy level, in the form of a photon, there is no energy shift (known as elastic or Rayleigh scattering), whereas when they return to a lower or a higher energy level there is a gain or loss of energy, respectively (known as inelastic or Raman scattering) [13]. The shift in the energy allows the generation of a spectrum which is indicative of the chemical bonds present in the sample. The characteristic spectra that are derived from Raman spectroscopy, represent a number of different biomolecules within a sample (*e.g.*, proteins, carbohydrates, lipids, DNA) [14]. Recent studies have employed Raman spectroscopy to study different diseases, such as malaria, oral and colorectal cancer, in biological fluids [15-17].

The aim of the present study was to diagnose patients with Alzheimer's disease, in early and late disease stages, and patients with DLB, as well as to discriminate between AD and DLB. To achieve this, blood plasma was analysed with Raman spectroscopy as a minimally invasive procedure that would also allow repeated measurements for follow-up of individuals.

Results

We enrolled 56 individuals into this study who were classified into 4 groups; early stage AD (n=11; age range: 50-74 years), late stage AD (n=15; age range: 50-79 years), DLB (n=15; age range: 23-73 years) and healthy controls (n=15; age range: 23-73 years) (Table 1). Early and late-stage AD was defined according to the duration of illness, from designated age at onset up to age at sample collection. P-values were calculated based on age and statistical differences were detected only for the following two subgroup comparisons: Late AD *vs* Healthy (P=0.004) and DLB *vs* Healthy (P <0.001). For all the other comparison groups (Early AD *vs* Healthy, Early AD *vs* DLB, Late AD *vs* DLB, Early AD *vs* Late AD), there was no statistical difference

observed due to age (P >0.005) (Supplementary Table 1). Even though there was age difference between the controls and AD individuals, no correlation was observed between age and AD spectra after using partial least squares regression (R^2 = 0.107, 2 latent variables with 99.93% cumulative explained variance) and no statistical difference was observed in the spectra of AD patients with age lower and higher than 54 years of age (average control age) with a 95% confidence level (P> 0.005). This indicates that age did not affect the spectral distribution within the AD class. Similarly, no statistical differences were observed in the Raman spectra of the different groups due to gender (male *vs* female) (Supplementary Fig. 8).

Early stage AD vs healthy individuals. After pre-processing of the spectral data, principal component analysis followed by linear discriminant analysis (PCA-LDA) was applied to the derived dataset. A one-dimensional (1D) scores plot was generated to account for differences and similarities between early stage AD and healthy subjects (Fig. 1A); after statistical analysis, the two classes showed significant differences (P <0.0001, 95% CI = 0.0503 to 0.0622). A loadings plot served as a biomarker extraction method, identifying the top six peaks responsible for differentiation: 1650 cm^{-1}, 1529 cm^{-1}, 1432 cm^{-1}, 1161 cm^{-1}, 996 cm^{-1} and 911 cm^{-1} (Fig. 1B). A statistical test was performed on each peak individually to calculate the P-value and investigate the differences in Raman intensity between the two groups (Supplementary Fig. S2, Supplementary Table 2). Figure 1C summarises the tentative assignments along with the P-values for these peaks (denoted with asterisks). Further analysis was conducted to classify the two classes; support vector machine (SVM) was the classification algorithm that was used, achieving 84% sensitivity and 86% specificity, with the G-score estimated at 85% and Youden's index at 70% (Table 2, Supplementary Fig. 1A).

Late AD vs healthy individuals. A similar approach was followed for the discrimination between late AD and healthy individuals. Figure 2A represents the scores plot after cross validated PCA-LDA and reveals statistically significant differences between the groups (P

<0.0001, 95% CI = 0.0655 to 0.0834). The top six discriminatory peaks that were selected for this comparison group were 1648 cm^{-1}, 1530 cm^{-1}, 1432 cm^{-1}, 1259 cm^{-1}, 1164 cm^{-1} and 1003 cm^{-1} (Fig. 2, Supplementary Fig. S3 and Supplementary Table 2). After the SVM classification, late AD was discriminated from healthy individuals with 84% sensitivity, 77% specificity and the G-score estimated at 80% and Youden's index at 61% (Table 2, Supplementary Fig. 1B).

DLB vs healthy individuals. Scores plot was again generated after cross validated PCA-LDA to compare DLB with healthy controls. Statistically significant differences were found between the groups (P <0.0001, 95% CI = 0.0982 to 0.1166) and the wavenumbers that were mostly responsible for this discrimination are shown in the respective loadings plot (Fig. 3B): 1647 cm^{-1}, 1604 cm^{-1}, 1418 cm^{-1}, 1384 cm^{-1}, 1002 cm^{-1} and 933 cm^{-1}. The differences in Raman intensity for each wavenumber are shown in more detail in Supplementary Fig. 4 and Supplementary Table 2. Sensitivity and specificity, after SVM, were 83% and 87%, respectively, while the G-score was calculated at 85% and Youden's index at 70% (Table 2, Supplementary Fig. 1C).

Early stage AD vs DLB. The scores plot for the comparison between early stage AD and DLB is shown in Fig. 4A. After statistical analysis, the difference between these two cohorts was statistically significant (P <0.0001, 95% CI = -0.0791 to -0.0649). The wavenumbers that were found as the most important, after cross-validated PCA-LDA are shown along with their tentative assignments in Fig. 4 and were the following: 1645 cm^{-1}, 1513 cm^{-1}, 1376 cm^{-1}, 1253 cm^{-1}, 1161 cm^{-1} and 1003 cm^{-1} (Supplementary Fig. 5 and Supplementary Table 2). The sensitivity and specificity values from this comparison were 81% and 88%, respectively, with the G-score at 84% and Youden's index at 69% (Table 2, Supplementary Fig. 1D).

Late AD vs DLB. Analyses were conducted to discriminate between late AD and DLB (Fig. 5). Significant differences were found after statistical analysis on the PCA-LDA scores

plot (P <0.0001; 95% CI = 0.138 to 0.1596). The following are the top six wavenumbers that were found to be responsible for the observed differentiation: 1646 cm^{-1}, 1614 cm^{-1}, 1437 cm^{-1}, 1216 cm^{-1}, 1164 cm^{-1} and 1003 cm^{-1}. Differences in the Raman intensity at these peaks are given in Supplementary Fig. 6 and Supplementary Table 2. The tentative assignments for these wavenumbers are shown in Fig. 5C. Sensitivity and specificity were 90% and 93%, respectively, with G-score being 91% and Youden's index at 84% (Table 2, Supplementary Fig. 1E).

Early AD vs late AD. A comparison between early and late-stage AD patients was also performed. After cross validated PCA-LDA, the scores plot revealed statistically significant differences between the two groups (P <0.0001; 95% CI = -0.0943 to -0.0624) (Fig. 6). The loadings plot denoted the following six wavenumbers as the most important: 1650 cm^{-1}, 1476 cm^{-1}, 1432 cm^{-1}, 1161 cm^{-1}, 1003 cm^{-1}, 642 cm^{-1} (Supplementary Fig. 7 and Supplementary Table 2). After classification of the two populations, 66% of the early AD spectra were correctly identified with 34% been misclassified as late AD; and 83% of the late AD cases were correctly identified with 17% misclassified as early stage AD (Table 2). G-score was calculated at 74% and Youden's index at 49% (Table 2, Supplementary Fig. 1F).

Discussion

Amyloid PET imaging has been shown to improve the diagnostic accuracy of AD [18]. However, one of the limitations is that only subjects with advanced dementia and relatively heavy plaque densities will be amyloid PET-positive; thus, individuals may not be identified early enough to be used in prevention studies using anti-amyloid therapeutics [19]. The detection accuracy of neuropathologically defined AD with PET imaging has been estimated at 69-95% sensitivity and 83-89% specificity [20]. In the case of DLB patients, PET imaging shows increased Aβ deposition in >50% of patients with DLB which limits its value in distinguishing between AD and DLB [3]. In a recent study, clinical and pathological diagnoses were compared

and DLB patients were identified with 73% sensitivity and 93% specificity; such findings suggest that there is still need for improvement in discriminating between these conditions [21]. When using MRI for AD diagnosis, a decreased volume of hippocampus and other temporal lobe structures is indicative of neurodegeneration; visual rating scales evaluating the degree of atrophy provide ~80-85% sensitivity and specificity when comparing AD to healthy individuals and slightly lower sensitivity and specificity when comparing to amnestic mild cognitive impairment (MCI) [6]. However, atrophy patterns can be similar in different diseases while at the same time some unusual forms of AD may have atypical patterns [22].

Established CSF biomarkers that are currently used in clinical practise to diagnose AD, also known as "core biomarkers", include decreased levels of Aβ42, or decreased Aβ42:Aβ40 ratio, and increased levels of total tau (T-tau) or hyperphosphorylated tau (P-tau) [23]. In a systematic review and meta-analysis, a number of different biomarkers has been associated with AD in both CSF and blood; namely, neurofilament light chain (NfL), neuron-specific enolase (NSE), visinin-like protein 1 (VLP-1), heart fatty acid binding protein (HFABP), chitinase-3-like protein 1 (YKL-40) in CSF, as well as T-tau and P-tau in blood plasma [10, 24, 25]. More recently, an elevated level of plasma NfL has been suggested as a promising biomarker to distinguish AD and MCI from healthy subjects. The accuracy for the comparison between AD and healthy controls, after testing for NfL, was 87%, which is comparable to accuracies achieved by CSF testing (88% Aβ42; 90% T-tau; 87% P-tau; 89% NfL) and plasma tau (78%) [11]. Another study, discovered and validated a set of ten lipids in plasma to detect preclinical AD in cognitively normal older adults within a 2-3 year timeframe; this panel achieved 90% accuracy [9]. Even though it is now established that the α-synuclein gene (*SNCA*) is associated with a few families with Parkinson's disease (PD) and DLB, CSF α-synuclein is not yet proven as a potential biomarker. CSF and blood biomarkers for the diagnosis of DLB remain elusive,

with Aβ, T-tau and P-tau remaining the most current measurements to predict cognitive decline and determine associated AD pathology [3].

In the present study, we included patients with AD, in both early and later stages of the disease, DLB, as well as heathy individuals. The blood-based Raman spectroscopic technique, provided excellent diagnostic accuracy not only between diseased and non-diseased states, but also between the two different types of dementia. Statistically significant age differences were only observed for Late AD vs Healthy (P=0.004) and DLB vs Healthy (P <0.001). The age difference between healthy controls and both Late AD and DLB patients was somehow expected as these diseases manifest mainly in older individuals. A larger dataset containing a wider age range would be necessary for adjusting the model for age. However, the fact that diagnostic accuracies remain exceptionally high for the subgroups with no age differences (*e.g.*, Late AD vs DLB showing 90% sensitivity and 93% specificity), implies that the age factor was not solely responsible for the achieved segregation between the cohorts. Similarly, no statistical differences were observed due to gender after calculating the P values for each spectral wavenumber; therefore, gender differences did not change the spectral profile.

Raman spectroscopy can reveal invaluable information about a biological sample as it provides the overall status of a sample, indicating disease. The results from such an approach are comparable to, and in some cases even better than, conventional methods, as they allow for simultaneous investigation of a panel of different biomarkers and therefore may be more suitable for complex diseases. Furthermore, Raman allows for a low-cost, label-free and non-destructive diagnosis in contrast to current imaging techniques and molecular CSF and/or blood tests. Previous studies have estimated the cost of an MRI and PET scan at £163 and £844, respectively, while an enzyme-linked immunosorbent assay (ELISA) measurement (96-well plate) of the core biomarkers (Aβ42, T-tau, P-tau) costs £826 per kit [26, 27]. In contrast, a blood test employing Raman spectroscopy is negligible in terms of consumables although there

would be costs in terms of employee time for samples preparation and analysis; overall cost would fall dramatically as the data infrastructure to allow remote classification of samples became available. Even the upfront cost or Raman instrumentation, often varying from £3,000-£150,000, is low in comparison with other approaches and would again fall with the development of hand-held devices; also the running costs are minimal with electrical power being the only requirement. Over the longer-term, lasers may need to be replaced (~every 6-7 years), but daily running costs are close to zero.

Discriminatory peaks have also been identified for all of the different comparison groups and could possibly be used as biomarkers for differential diagnosis or screening of high-risk populations. For instance, higher levels of Amide II peaks (~1530 cm^{-1}) were seen in both early ($P < 0.0001$) and late stage AD ($P < 0.0001$) patients and could possibly be represented by an increase in tau proteins or NfL in plasma, which have been suggested previously as promising biomarkers (Supplementary Fig. 2, Supplementary Fig. 3) [11]. Also, the observed decrease in lipids (~1432 cm^{-1}) could be due to damaged phospholipid membranes caused by oxidative stress. These findings are in line with previous results of a larger-scale study our research team conducted, in which infrared (IR) spectroscopy was employed to diagnose AD [28]. An advantage of Raman spectroscopy over IR is its ability to analyse aqueous samples which would allow the analysis of fresh samples without the need of prior dehydration; this would be particularly beneficial for use in a clinic. Noticeably, in this preceding study, lipid peaks were also decreased (~1740 cm^{-1}, $P < 0.05$; ~1450 cm^{-1}, $P < 0.005$) and Amide II was also increased (~1540 cm^{-1}, $P = 0.003$) in AD patients. However, Amide I (~1650 cm^{-1}), which is indicative of Aβ load, was not found to be statistically different ($P = 0.12$), in contrast to the current study where it was significant in both early ($P = 0.0003$, 95% CI = 0.0008 to 0.0028) and late stage AD ($P < 0.0001$, 95% CI = 0.0016 to 0.0029). Previous studies have noted altered levels of aromatic amino acids in plasma and serum of AD patients [29]. Some studies have shown an

increase in phenylalanine in the brain of AD subjects [30-32], while others suggest a decrease [33, 34]. In our study, the level of phenylalanine was increased in DLB cases, whereas in late AD phenylalanine was decreased when compared to healthy subjects. Between AD and DLB patients, the latter cohort showed higher levels of phenylalanine, which could possibly relate to their α-synuclein pathology (Supplementary Fig. 5). Previous studies have shown altered metabolic profiles of PD patients (also related to α-synuclein aggregation) when compared to normal controls, and these differences were related to metabolic pathway variations such as phenylalanine metabolism [35-37].

We were particularly interested in examining early stage AD cases as it is of crucial importance to identify individuals before brain damage becomes very severe. Evidence of changes here would allow for an on-time intervention, potentially to slow down the disease, psychologically prepare the affected person and their family, as well as provide them with the opportunity to take part in early intervention trials. Surprisingly, the diagnostic accuracy was slightly higher for early AD than for late AD. After comparison of these two groups, 66% of early AD and 83% of late AD were correctly classified. Of the wavelengths which were shown to contribute the most to the segregation between the classes, a peak assigned to Amide II proteins (~1476 cm^{-1}) and a peak assigned to C-C and C-S vibrations of proteins (~642 cm^{-1}) were found to be statistically significant (Supplementary Fig. 7). A potential explanation for the decreased level of Amide II in early stage AD cases could be the lower density of neurofibrillary tangles in the brain during early stages. Previous studies have suggested that kinase mutations and dysfunction play an important role in the development of disorders such as cancer and neurodegeneration [38]. Specifically, cyclin-dependent kinase 5 (cdk5), which is involved in the abnormal hyperphosphorylation of tau, has been suggested to accumulate at a relatively early stage in the neocortex [39]; more recent research has also shown that a cellular stress response, caused by accumulation of misfolded proteins, induces the activity of a major

tau kinase (GSK-3β) and occurs at an early stage of neurofibrillary degeneration leading to AD pathogenesis [40]. Therefore, this may potentially explain the increased level of the protein peak at 642 cm^{-1}. Special attention was also given to the accurate diagnosis of DLB and differentiation from AD which is especially important to provide the appropriate treatment; DLB cases respond well to cholinesterase inhibitors but have severe neuroleptic sensitivity reactions, which are associated with significantly increased morbidity and mortality [41].

A critical aspect for every new biomarker, diagnostic or treatment approach is the repetition and validation of the analytical process and in different cohorts. Previously, a few studies also employed Raman spectroscopy to diagnose AD in blood, achieving high classification accuracy. Carmona *et al*. distinguished AD (n=35) from normal (n=12) with 89% sensitivity and 92% specificity [42]. Ryzhikova *et al.* included serum samples from 20 AD patients, 18 patients with other neurodegenerative dementias (OD) (5 with DLB, 10 with Parkinson's disease dementia and 3 with frontotemporal dementia) and 10 healthy individuals and achieved 95% sensitivity and specificity [43]. However, the fact that a range of different dementias were all taken in the same group, may obscure the actual classification capability between AD and DLB. Moreover, no spectroscopic approach has been employed so far to investigate DLB in more detail.

A limitation of the current study is the small number of participants, which can affect sensitivity and specificity estimates. However, G-score values were estimated at 74-91%, denoting that the models were not overfitted. G-score does not account the size of classes, thus providing robust information about the classification ability even in smaller cohorts [44]. Youden's index values ranged between 49% (early AD *vs* late AD) and 84% (late AD *vs* DLB). This parameter is a probability indicator of the model's ability to avoid failure. Youden's indexes above 70% for early AD *vs* healthy, DLB *vs* healthy and late AD *vs* DLB indicate that these models have low probability of misclassification in the future. Another limitation of this

study, as well as similar previous ones, is the lack of serial samples from the same individuals which would validate the results and demonstrate repeatability.

In summary, diagnosis of early stage AD, late stage AD, DLB as well as differentiation between the two dementias was achieved, opening a new road for potential applications in a clinical setting. Some of the future uses of spectroscopy could be the detection of prodromal/pre-demented cases; the differential diagnosis of different dementias that would allow the appropriate treatment and/or recruitment into clinical trials; and the further monitoring of patients that do finally take part in clinical trials.

Methods

Patient information. We enrolled 56 individuals into this study who were classified into four groups; early stage AD (n=11), late stage AD (n=15), DLB (n=15) and healthy controls, usually spouses (n=15). Early and late-stage AD was defined according to the duration of illness, from designated age at onset up to age at sample collection. Early stage was defined as up to two years from designated age at onset, whereas late stage AD was defined as any duration beyond this time point. Clinical and demographic data is summarised in Table 1. Information on apolipoprotein ε4 (*APOE4)* status and gender was not available for two subjects from the healthy control group. Patients were recruited at Salford Royal Hospital (Salford, UK) with informed consent prior to enrolment in accordance with Local Ethical Approval (05/Q1405/24 conferred by North West 10 Research Ethics Committee Greater Manchester North). Patients were diagnosed according to battery of psychological testing (Manchester Neuropsychology Inventory) performed at a Specialist referral Centre (Cerebral Function Unit, Greater Manchester Neurosciences Centre, Salford Royal Hospital). All methods were performed in accordance with the relevant guidelines and all other applicable laws and regulations. At time of diagnosis patients were not receiving any medications, such as

anticholinesterase treatments. Most patients had received MRI scans but these were used only to support the neuropsychological outcomes.

Sample preparation and *APOE* genotyping. Whole blood samples were collected into EDTA tubes, centrifuged at 2000 rpm at 4^{o}C for 10 min to separate erythrocytes from plasma. Plasma was collected in 0.5 mL clean, plastic tubes, stored at -80^{o}C and thawed at room temperature prior to spectroscopic interrogation. After the samples were thawed, 50 μL were deposited on glass slides covered with aluminium foil, which has been shown to be featureless in Raman [14], and were then left to air-dry overnight. DNA was extracted by routine methods from blood samples of patients and control subjects; *APOE* alleles were determined by PCR [45].

Raman spectroscopy. Raman spectra were collected with an InVia Renishaw Raman spectrometer coupled with a charge-coupled device (CCD) detector and a Leica microscope. A 200 mW laser diode was used at a wavelength of 785 nm with a grating of 1200 l/mm, and the system was calibrated to 520.5 cm^{-1} with a silicon source, before every run. After trial-and-error measurements to optimise the experimental parameters, we concluded to a 10 second exposure time, 5% laser power and 2 accumulations at a spectral range 2000-400 cm^{-1} to achieve optimum spectral quality. Twenty-five point spectra were taken per sample using a 50× objective to focus the laser beam on the sample.

Pre-processing of spectral data and multivariate analysis. Spectra were initially corrected for cosmic rays using the Renishaw WiRe software. An in-house developed IRootLab toolbox (http://trevisanj.github.io/irootlab/) was then implemented within MATLAB (MathWorks, Natick, USA) for further pre-processing and computational analysis of the data. All spectra were cut at 1750-500 cm^{-1}, first order differentiated with Savitzky-Golay (SG) (window of 9 points; 2nd polynomial filter) to smooth out the noise and vector normalised to

account for non-biological differences, such as varying concentration or thickness of the sample; the resulting dataset was then mean-centered before implementation of cross-validated (k-fold=10, leave-one-out) principal component analysis followed by linear discriminant analysis (PCA-LDA). The leave-one-out cross-validation was implemented to avoid overfitting. This ensures that one sample is removed from the training set and predicted as external sample during model construction in an interactive process until all samples are predicted; this provides more realistic classification results. All classification models were validated using 10% of the samples in a test set. PCA is an unsupervised method that reduces the spectral dataset to only a few important principal components (PCs) which are responsible for the majority of the variation; using a Pareto function, a number of 10 PCs was found as optimal. LDA is a supervised technique, often coupled with PCA, to maximise the between-class distance and minimise the within-class distance. Scores plots and loadings plots were generated after PCA-LDA to visualise the differences and similarities between the groups as well as to identify specific spectral peaks responsible for this differentiation; these peaks were tentatively assigned to different biomolecules which can potentially serve as biomarkers [46, 47]. After the six peaks were identified from the loadings plot, they were then extracted from polynomial corrected, vector normalised spectra in order to avoid the spectral transformation that first order differentiation can cause. Classification of the different comparison groups was conducted using support vector machine (SVM) which is a machine-learning technique to classify spectral data. For SVM implementation, the pre-processed dataset (*i.e.*, cut, SG differentiated, vector normalised) was normalised to the [0, 1] range and then the optimal (C, γ) combination was found using grid search. Sensitivities and specificities were therefore calculated for each comparison group [48]. In order to overcome the limitation of using a small cohort in this study, G-score values were also calculated to assess the overall performance of the classification model [44]. The G-score is calculated as the square root of sensitivity times

specificity. Youden's index was calculated to assess the classifier's ability to avoid failure. This parameter is estimated as sensitivity minus (1 – specificity).

Statistical analysis. The values generated after cross-validated PCA-LDA, were imported into GraphPad Prism 7 to conduct the statistical analyses and calculate the P-values for each comparison. Differences between two groups were assessed using a Student's t-test (two-tailed, non-parametric, Mann-Whitney test, 95% confidence interval). The data were expressed as the mean ± standard deviation (SD). A P-value of 0.05 or less was considered significant in all statistical tests.

Additional Information

Availability of data and material

All data (raw and pre-processed spectra) along with appropriate code identifiers will be uploaded onto the publicly accessible data repository Figshare.

Conflict of Interest Disclosure

The authors declare that they have no competing interests.

Funding

This study was supported by Rosemere Cancer Foundation (RCF).

Authors' contributions

M.P performed the experiments, analysed the data and wrote the manuscript. The manuscript was written with contributions from C.L.M.M, D.E.H, D.M.A.M, D.A, P.L.M-H and F.L.M. All authors have read and approved the final manuscript.

Acknowledgements

MP would like to thank the Rosemere Cancer Foundation (RCF) for funding. CLMM would like to thank CAPES-Brazil (grant 88881.128982/2016-01) for financial support.

Supplementary Information

Sensitivity and specificity rates after classification with support vector machine (SVM); differences in the intensity levels of important biomolecules after the comparison between early-stage AD and healthy individuals; differences in the intensity levels of important biomolecules after the comparison between late-stage AD and healthy individuals; differences in the intensity levels of important biomolecules after the comparison between DLB and healthy individuals; differences in the intensity levels of important biomolecules after the comparison between early-stage AD and DLB; differences in the intensity levels of important biomolecules after the comparison between late-stage AD and DLB; differences in the intensity levels of important biomolecules after the comparison between early-stage AD and late-stage AD; mean values, standard deviations (SD), 95% confidence intervals (CI) and P-values after statistical analysis for the different comparison groups; P-values (two-tail, 95% confidence level) for the different ages in each comparison group.

References

[1] Minoshima, S., Foster, N. L., Sima, A. A., Frey, K. A., Albin, R. L., and Kuhl, D. E. (2001) Alzheimer's disease versus dementia with Lewy bodies: cerebral metabolic distinction with autopsy confirmation, *Ann Neurol 50*, 358-365.

[2] Mueller, C., Ballard, C., Corbett, A., and Aarsland, D. (2017) The prognosis of dementia with Lewy bodies, *Lancet Neurol*.

[3] McKeith, I. G., Boeve, B. F., Dickson, D. W., Halliday, G., Taylor, J.-P., Weintraub, D., Aarsland, D., Galvin, J., Attems, J., Ballard, C. G., Bayston, A., Beach, T. G., Blanc, F., Bohnen, N., Bonanni, L., Bras, J., Brundin, P., Burn, D., Chen-Plotkin, A., Duda, J. E., El-Agnaf, O., Feldman, H., Ferman, T. J., ffytche, D., Fujishiro, H., Galasko, D., Goldman, J. G., Gomperts, S. N., Graff-Radford, N. R., Honig, L. S., Iranzo, A., Kantarci, K., Kaufer, D., Kukull, W., Lee, V. M. Y., Leverenz, J. B., Lewis, S., Lippa, C., Lunde, A., Masellis, M., Masliah, E., McLean, P., Mollenhauer, B., Montine, T. J., Moreno, E., Mori, E., Murray, M., O'Brien, J. T., Orimo, S., Postuma, R. B., Ramaswamy, S., Ross, O. A., Salmon, D. P., Singleton, A., Taylor, A., Thomas, A., Tiraboschi, P., Toledo, J. B., Trojanowski, J. Q., Tsuang, D., Walker, Z., Yamada, M., and Kosaka, K. (2017) Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB Consortium, *Neurology 89*, 88-100.

[4] Walker, Z., Possin, K. L., Boeve, B. F., and Aarsland, D. (2015) Lewy body dementias, *Lancet 386*, 1683-1697.

[5] McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., Mohs, R. C., Morris, J. C., Rossor, M. N., Scheltens, P., Carrillo, M. C., Thies, B., Weintraub, S., and Phelps, C. H. (2011) The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease, *Alzheimers Dement 7*, 263-269.

[6] Frisoni, G. B., Fox, N. C., Jack, C. R., Scheltens, P., and Thompson, P. M. (2010) The clinical use of structural MRI in Alzheimer disease, *Nat Rev Neurol 6*, 67-77.

[7] Saint-Aubert, L., Lemoine, L., Chiotis, K., Leuzy, A., Rodriguez-Vieitez, E., and Nordberg, A. (2017) Tau PET imaging: present and future directions, *Mol Neurodegener 12*, 19.

[8] Ossenkoppele, R., Schonhaut, D. R., Schöll, M., Lockhart, S. N., Ayakta, N., Baker, S. L., O'Neil, J. P., Janabi, M., Lazaris, A., and Cantwell, A. (2016) Tau PET patterns mirror clinical and neuroanatomical variability in Alzheimer's disease, *Brain 139*, 1551-1567.

[9] Mapstone, M., Cheema, A. K., Fiandaca, M. S., Zhong, X., Mhyre, T. R., MacArthur, L. H., Hall, W. J., Fisher, S. G., Peterson, D. R., Haley, J. M., Nazar, M. D., Rich, S. A., Berlau, D. J., Peltz, C. B.,

Tan, M. T., Kawas, C. H., and Federoff, H. J. (2014) Plasma phospholipids identify antecedent memory impairment in older adults, *Nat Med 20*, 415-418.
[10] Olsson, B., Lautner, R., Andreasson, U., Öhrfelt, A., Portelius, E., Bjerke, M., Hölttä, M., Rosén, C., Olsson, C., and Strobel, G. (2016) CSF and blood biomarkers for the diagnosis of Alzheimer's disease: a systematic review and meta-analysis, *Lancet Neurol 15*, 673-684.
[11] Mattsson, N., Andreasson, U., Zetterberg, H., Blennow, K., and for the Alzheimer's Disease Neuroimaging, I. (2017) Association of plasma neurofilament light with neurodegeneration in patients with alzheimer disease, *JAMA Neurol 74*, 557-566.
[12] Hye, A., Lynham, S., Thambisetty, M., Causevic, M., Campbell, J., Byers, H. L., Hooper, C., Rijsdijk, F., Tabrizi, S. J., Banner, S., Shaw, C. E., Foy, C., Poppe, M., Archer, N., Hamilton, G., Powell, J., Brown, R. G., Sham, P., Ward, M., and Lovestone, S. (2006) Proteome-based plasma biomarkers for Alzheimer's disease, *Brain 129*, 3042-3050.
[13] Raman, C. V., and Krishnan, K. S. (1928) A new type of secondary radiation, *Nature 121*, 501-502.
[14] Butler, H. J., Ashton, L., Bird, B., Cinque, G., Curtis, K., Dorney, J., Esmonde-White, K., Fullwood, N. J., Gardner, B., Martin-Hirsch, P. L., Walsh, M. J., McAinsh, M. R., Stone, N., and Martin, F. L. (2016) Using Raman spectroscopy to characterize biological materials, *Nat Protoc 11*, 664-687.
[15] Sahu, A., Nandakumar, N., Sawant, S., and Krishna, C. M. (2015) Recurrence prediction in oral cancers: a serum Raman spectroscopy study, *Analyst 140*, 2294-2301.
[16] Li, X., Yang, T., and Li, S. (2012) Discrimination of serum Raman spectroscopy between normal and colorectal cancer using selected parameters and regression-discriminant analysis, *Phys Chem Chem Phys 51*, 5038-5043.
[17] Garrett, N. L., Sekine, R., Dixon, M. W., Tilley, L., Bambery, K. R., and Wood, B. R. (2015) Bio-sensing with butterfly wings: naturally occurring nano-structures for SERS-based malaria parasite detection, *Phys Chem Chem Phys 17*, 21164-21168.
[18] Foster, N. L., Heidebrink, J. L., Clark, C. M., Jagust, W. J., Arnold, S. E., Barbas, N. R., DeCarli, C. S., Scott Turner, R., Koeppe, R. A., Higdon, R., and Minoshima, S. (2007) FDG-PET improves accuracy in distinguishing frontotemporal dementia and Alzheimer's disease, *Brain 130*, 2616-2635.
[19] Beach, T. G. (2017) A Review of Biomarkers for Neurodegenerative Disease: Will They Swing Us Across the Valley?, *Neurol Ther 6*, 5-13.
[20] Beach, T. G., Schneider, J. A., Sue, L. I., Serrano, G., Dugger, B. N., Monsell, S. E., and Kukull, W. (2014) Theoretical impact of Florbetapir (18F) amyloid imaging on diagnosis of alzheimer dementia and detection of preclinical cortical amyloid, *J Neuropathol Exp Neurol 73*, 948-953.
[21] Skogseth, R. E., Hortobágyi, T., Soennesyn, H., Chwiszczuk, L., Rongve, A., Ballard, C., and Aarsland, D. (2017) Accuracy of Clinical Diagnosis of Dementia with Lewy Bodies versus Neuropathology, *J Alzheimers Dis*, 1-14.
[22] Johnson, K. A., Fox, N. C., Sperling, R. A., and Klunk, W. E. (2012) Brain Imaging in Alzheimer Disease, *Cold Spring Harb Perspect Med 2*, a006213.
[23] Frisoni, G. B., Boccardi, M., Barkhof, F., Blennow, K., Cappa, S., Chiotis, K., Démonet, J.-F., Garibotto, V., Giannakopoulos, P., and Gietl, A. (2017) Strategic roadmap for an early diagnosis of Alzheimer's disease based on biomarkers, *Lancet Neurol 16*, 661-676.
[24] Zetterberg, H., Wilson, D., Andreasson, U., Minthon, L., Blennow, K., Randall, J., and Hansson, O. (2013) Plasma tau levels in Alzheimer's disease, *Alzheimers Res Ther 5*, 9.
[25] Tatebe, H., Kasai, T., Ohmichi, T., Kishi, Y., Kakeya, T., Waragai, M., Kondo, M., Allsop, D., and Tokuda, T. (2017) Quantification of plasma phosphorylated tau to use as a biomarker for brain Alzheimer pathology: pilot case-control studies including patients with Alzheimer's disease and down syndrome, *Mol Neurodegener 12*, 63.
[26] Humpel, C. (2011) Identifying and validating biomarkers for Alzheimer's disease, *Trends Biotechnol 29*, 26-32.

[27] Escudero, J., Ifeachor, E., Zajicek, J. P., Green, C., Shearer, J., and Pearson, S. (2013) Machine learning-based method for personalized and cost-effective detection of Alzheimer's disease, *IEEE Trans Biomed Eng 60*, 164-168.
[28] Paraskevaidi, M., Morais, C. L., Lima, K. M., Snowden, J. S., Saxon, J. A., Richardson, A. M., Jones, M., Mann, D. M., Allsop, D., and Martin-Hirsch, P. L. (2017) Differential diagnosis of Alzheimer's disease using spectrochemical analysis of blood, *Proc Natl Acad Sci USA*, 201701517.
[29] Griffin, J. W., and Bradshaw, P. C. (2017) Amino Acid Catabolism in Alzheimer's Disease Brain: Friend or Foe?, *Oxid Med Cell Longev 2017*.
[30] Xu, J., Begley, P., Church, S. J., Patassini, S., Hollywood, K. A., Jullig, M., Curtis, M. A., Waldvogel, H. J., Faull, R. L., Unwin, R. D., and Cooper, G. J. (2016) Graded perturbations of metabolism in multiple regions of human brain in Alzheimer's disease: Snapshot of a pervasive metabolic disorder, *Biochim Biophys Acta 1862*, 1084-1092.
[31] Nilsen, L. H., Witter, M. P., and Sonnewald, U. (2014) Neuronal and astrocytic metabolism in a transgenic rat model of Alzheimer's disease, *J Cereb Blood Flow Metab 34*, 906-914.
[32] Wissmann, P., Geisler, S., Leblhuber, F., and Fuchs, D. (2013) Immune activation in patients with Alzheimer's disease is associated with high serum phenylalanine concentrations, *J Neurol Sci 329*, 29-33.
[33] Gonzalez-Dominguez, R., Garcia-Barrera, T., and Gomez-Ariza, J. L. (2015) Metabolite profiling for the identification of altered metabolic pathways in Alzheimer's disease, *J Pharm Biomed Anal 107*, 75-81.
[34] Trushina, E., Dutta, T., Persson, X.-M. T., Mielke, M. M., and Petersen, R. C. (2013) Identification of altered metabolic pathways in plasma and CSF in mild cognitive impairment and Alzheimer's disease using metabolomics, *PloS one 8*, e63644.
[35] Luan, H., Liu, L. F., Meng, N., Tang, Z., Chua, K. K., Chen, L. L., Song, J. X., Mok, V. C., Xie, L. X., Li, M., and Cai, Z. (2015) LC-MS-based urinary metabolite signatures in idiopathic Parkinson's disease, *J Proteome Res 14*, 467-478.
[36] Luan, H., Liu, L.-F., Tang, Z., Zhang, M., Chua, K.-K., Song, J.-X., Mok, V. C., Li, M., and Cai, Z. (2015) Comprehensive urinary metabolomic profiling and identification of potential noninvasive marker for idiopathic Parkinson's disease, *Sci Rep 5*.
[37] Havelund, J. F., Heegaard, N. H., Færgeman, N. J., and Gramsbergen, J. B. (2017) Biomarker Research in Parkinson's Disease Using Metabolite Profiling, *Metabolites 7*, 42.
[38] Lahiry, P., Torkamani, A., Schork, N. J., and Hegele, R. A. (2010) Kinase mutations in human disease: interpreting genotype-phenotype relationships, *Nat Rev Genet 11*, 60.
[39] Pei, J.-J., Grundke-Iqbal, I., Iqbal, K., Bogdanovic, N., Winblad, B., and Cowburn, R. F. (1998) Accumulation of cyclin-dependent kinase 5 (cdk5) in neurons with early stages of Alzheimer's disease neurofibrillary degeneration, *Brain Res 797*, 267-277.
[40] Hoozemans, J. J., van Haastert, E. S., Nijholt, D. A., Rozemuller, A. J., Eikelenboom, P., and Scheper, W. (2009) The unfolded protein response is activated in pretangle neurons in Alzheimer's disease hippocampus, *Am J Pathol 174*, 1241-1251.
[41] McKeith, I. (2004) Dementia with Lewy bodies, *Dialogues Clin Neurosci 6*, 333-341.
[42] Carmona, P., Molina, M., Calero, M., Bermejo-Pareja, F., Martinez-Martin, P., and Toledano, A. (2013) Discrimination analysis of blood plasma associated with Alzheimer's disease using vibrational spectroscopy, *J Alzheimers Dis 34*, 911-920.
[43] Ryzhikova, E., Kazakov, O., Halamkova, L., Celmins, D., Malone, P., Molho, E., Zimmerman, E. A., and Lednev, I. K. (2015) Raman spectroscopy of blood serum for Alzheimer's disease diagnostics: specificity relative to other types of dementia, *J Biophotonics 8*, 584-596.
[44] Parikh, K. S., and Shah, T. P. (2016) Support vector machine–a large margin classifier to diagnose skin illnesses, *Procedia Technology 23*, 369-375.

[45] Davidson, Y., Gibbons, L., Purandare, N., Byrne, J., Hardicre, J., Wren, J., Payton, A., Pendleton, N., Horan, M., Burns, A., and Mann, D. M. A. (2006) Apolipoprotein E υ4 Allele Frequency in Vascular Dementia, *Dement Geriatr Cogn Dis 22*, 15-19.

[46] Movasaghi, Z., Rehman, S., and Rehman, I. U. (2007) Raman spectroscopy of biological tissues, *Appl Spectrosc Rev 42*, 493-541.

[47] Pichardo-Molina, J. L., Frausto-Reyes, C., Barbosa-García, O., Huerta-Franco, R., González-Trujillo, J. L., Ramírez-Alvarado, C. A., Gutiérrez-Juárez, G., and Medina-Gutiérrez, C. (2007) Raman spectroscopy and multivariate analysis of serum samples from breast cancer patients, *Lasers Med Sci 22*, 229-236.

[48] Trevisan, J., Angelov, P. P., Carmichael, P. L., Scott, A. D., and Martin, F. L. (2012) Extracting biological information with computational analysis of Fourier-transform infrared (FTIR) biospectroscopy datasets: current practices to future perspectives, *Analyst 137*, 3202-3215.

Legends for the figures

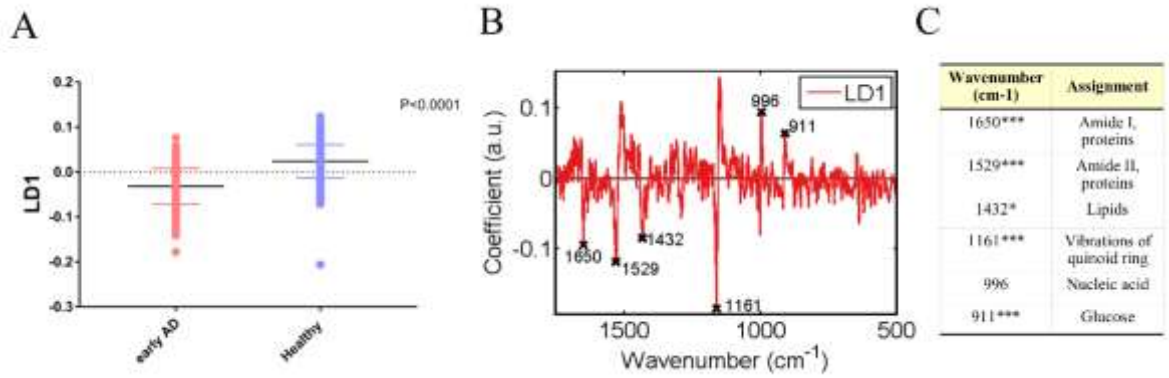
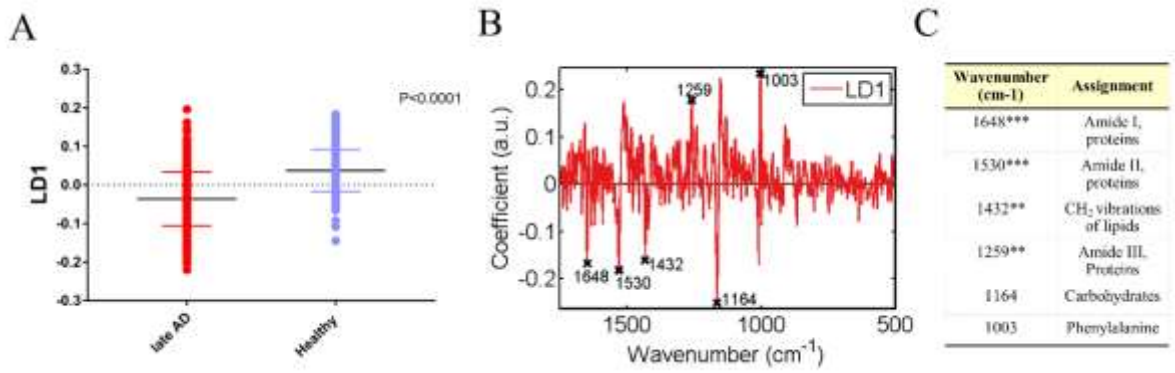
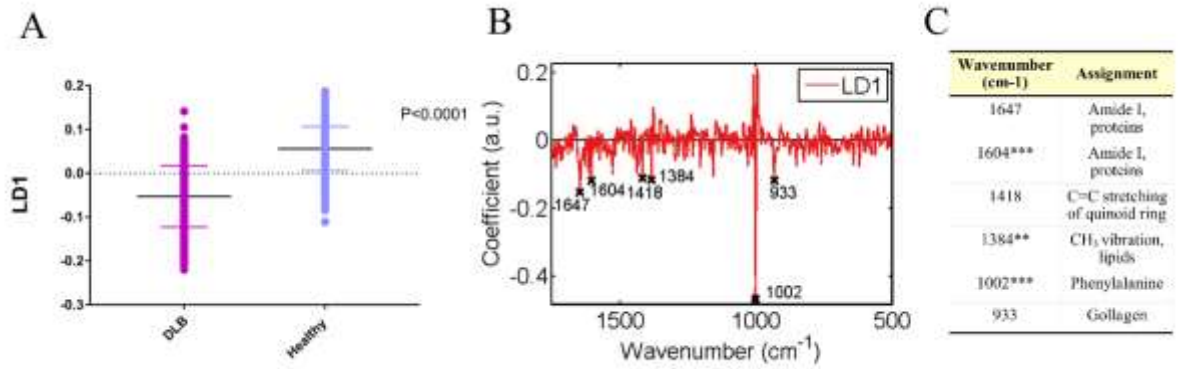


Figure 1. Early stage Alzheimer's disease (AD) *versus* healthy individuals. One-dimensional (1D) scores plot after cross-validated PCA-LDA ($P < 0.0001$, 95% CI = 0.0503 to 0.0622) (A); loadings plot showing the top six discriminatory peaks between the two classes (B); important peaks along with their tentative assignments [45,46] (C). Data are expressed as the mean ± standard deviation (SD). A P-value of 0.05 or less was considered significant; $P < 0.05$ (*) or $P < 0.005$ (**) or $P < 0.0005$ (***).



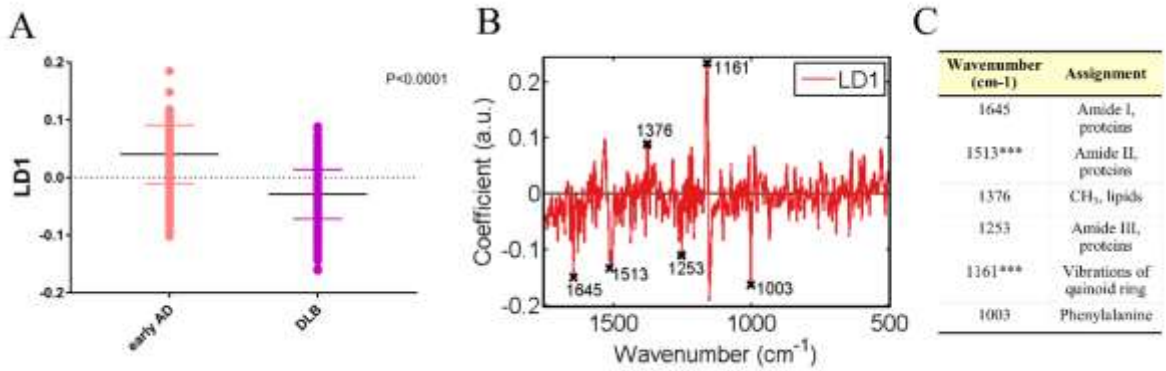
Wavenumber (cm-1)	Assignment
1648***	Amide I, proteins
1530***	Amide II, proteins
1432**	CH_2 vibrations of lipids
1259**	Amide III, Proteins
1164	Carbohydrates
1003	Phenylalanine

Figure 2. Late stage Alzheimer's disease (AD) *versus* healthy individuals. One-dimensional (1D) scores plot after cross-validated PCA-LDA (P <0.0001, 95% CI = 0.0655 to 0.0834) (A); loadings plot showing the top six discriminatory peaks between the two classes (B); important peaks along with their tentative assignments [45,46] (C). Data are expressed as the mean ± standard deviation (SD). A P-value of 0.05 or less was considered significant; P <0.05 (*) or P <0.005 (**) or P <0.0005 (***).



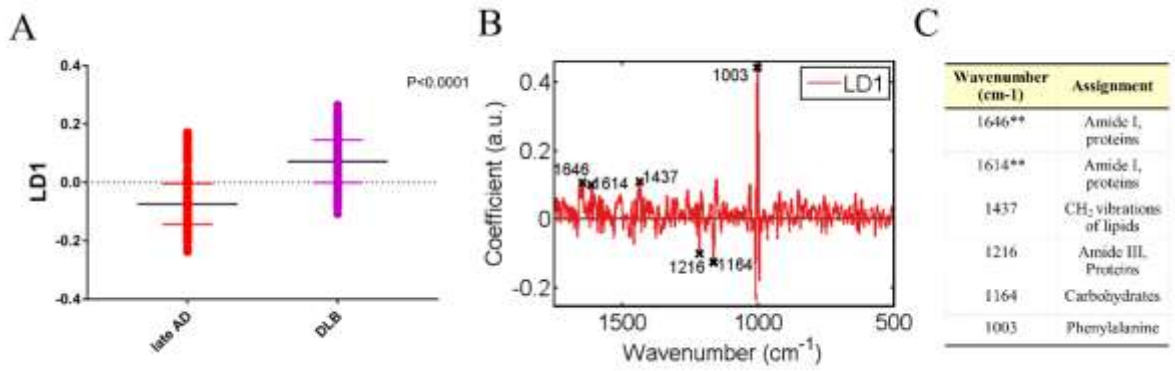
Wavenumber (cm-1)	Assignment
1647	Amide I, proteins
1604***	Amide I, proteins
1418	C=C stretching of quinoid ring
1384**	CH_3 vibration, lipids
1002***	Phenylalanine
933	Collagen

Figure 3. Dementia with Lewy bodies (DLB) *versus* healthy individuals. One-dimensional (1D) scores plot after cross-validated PCA-LDA (P <0.0001, 95% CI = 0.0982 to 0.1166) (A); loadings plot showing the top six discriminatory peaks (B); important peaks along with their tentative assignments [45,46] (C). Data are expressed as the mean ± standard deviation (SD). A P-value of 0.05 or less was considered significant; P <0.05 (*) or P <0.005 (**) or P <0.0005 (***).



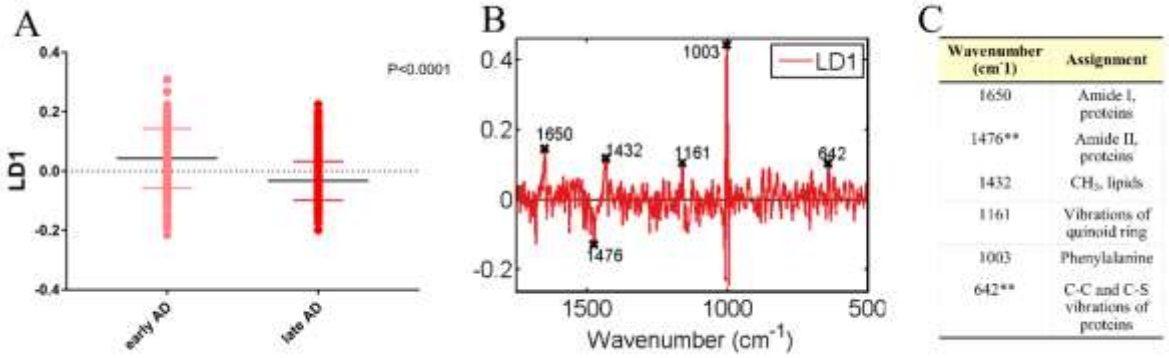
Wavenumber (cm-1)	Assignment
1645	Amide I, proteins
1513***	Amide II, proteins
1376	CH_3, lipids
1253	Amide III, proteins
1161***	Vibrations of quinoid ring
1003	Phenylalanine

Figure 4. Early stage Alzheimer's disease (AD) *versus* dementia with Lewy bodies (DLB). One-dimensional (1D) scores plot after cross-validated PCA-LDA (P <0.0001, 95% CI = -0.0791 to -0.0649) (A); loadings plot showing the top six discriminatory peaks (early AD was used as reference class) between the two classes (B); important peaks along with their tentative assignments [45,46] (C). Data are expressed as the mean ± standard deviation (SD). A P-value of 0.05 or less was considered significant; P <0.05 (*) or P <0.005 (**) or P <0.0005 (***).



Wavenumber (cm-1)	Assignment
1646**	Amide I, proteins
1614**	Amide I, proteins
1437	CH_2 vibrations of lipids
1216	Amide III, Proteins
1164	Carbohydrates
1003	Phenylalanine

Figure 5. Late stage Alzheimer's disease (AD) *versus* dementia with Lewy bodies (DLB). One-dimensional (1D) scores plot after cross-validated PCA-LDA (P <0.0001, 95% CI = 0.138 to 0.1596) (A); loadings plot showing the top six discriminatory peaks (late AD was used as reference class) (B); important peaks along with their tentative assignments [45,46] (C). Data are expressed as the mean ± standard deviation (SD). A P-value of 0.05 or less was considered significant; P <0.05 (*) or P <0.005 (**) or P <0.0005 (***).



Wavenumber (cm-1)	Assignment
1650	Amide I, proteins
1476**	Amide II, proteins
1432	CH_3, lipids
1161	Vibrations of quinoid ring
1003	Phenylalanine
642**	C-C and C-S vibrations of proteins

Figure 6. Early stage Alzheimer's disease (AD) *versus* late AD. One-dimensional (1D) scores plot after cross-validated PCA-LDA (P <0.0001, 95% CI = -0.0943 to -0.0624) (A); loadings plot showing the top six discriminatory peaks (early AD was used as reference class) (B); important peaks along with their tentative assignments [45,46]. (C). Data are expressed as the mean ± standard deviation (SD). A P-value of 0.05 or less was considered significant; P <0.05 (*) or P <0.005 (**) or P <0.0005 (***).

Table 1. Patient characteristics.

	Early stage AD	Late stage AD	DLB	Healthy[a]
Sample size, n	11	15	15	15
Age, years, mean (SD, range)	62 (10, 50-74)	64 (8, 50-79)	71 (6, 61-80)	54 (18, 23-73)
***APOE4* carriers, n (%)**	6 (55)	11 (73)	6 (40)	6 (40)
Female, n (%)	5 (45)	3 (20)	3 (20)	9 (60)
Duration, years, mean (± SD)	1.28 (±0.5)	4.56 (±3)	2.46 (±1)	n/a

AD: Alzheimer's disease DLB: dementia with Lewy bodies; *APOE4:* apolipoprotein E4; n/a: not applicable

[a] Two individuals from the 'Healthy' group had no information on *APOE4* load (13%) and gender (13%).

Table 2. Sensitivity, specificity and G-score and Youden's index for the different comparison groups after classification with support vector machine (SVM).

Comparison Group	Sensitivity (%)	Specificity (%)	G-Score (%)	Youden's index (%)
Early AD vs Healthy	84	86	85	70
Late AD vs Healthy	84	77	80	61
DLB vs Healthy	83	87	85	70
Early AD vs DLB	81	88	84	69
Late AD vs DLB	90	93	91	84
Early AD vs Late AD	66	83	74	49

AD: Alzheimer's disease; DLB: dementia with Lewy bodies