Identification of Gender Specific Biomarkers for Alzheimer's Disease

Dragan Gamberger[1(✉)], Bernard Ženko[2], Alexis Mitelpunkt[3], and Nada Lavrač[2,4] for the Alzheimer's Disease Neuroimaging Initiative

[1] Rudjer Bošković Institute, Bijenička 54, 10000 Zagreb, Croatia
dragan.gamberger@irb.hr
[2] Jožef Stefan Institute, Ljubljana, Slovenia
[3] Tel Aviv University, Tel Aviv, Israel
[4] University of Nova Gorica, Nova Gorica, Slovenia

Abstract. The paper presents experiments with a novel clustering methodology that enables identification of subpopulations of the Alzheimer's disease patients that are homogeneous in respect of both clinical and biological descriptors. It is expected that recognition of relevant connections between clinical and biological descriptors will be easier within such subpopulations. Our dataset includes 317 female and 342 male patients from the ADNI study that are described by a total of 243 biological and clinical descriptors recorded at baseline evaluation. The constructed clusters clearly demonstrate differences between female and male patient subpopulations. An interesting result is identification of a cluster of male Alzheimer's disease patients that are, surprisingly, characterized by increased intracerebral and whole brain volumes. The finding suggests existence of two different biological pathways for the Alzheimer's disease.

1 Introduction

Identification of connections between biological and clinical characteristics of Alzheimer's disease patients is a long term goal that could significantly improve the understanding of the Alzheimer's disease (AD) pathophysiology, improve clinical trial design, and help in predicting outcomes of mild cognitive impairment [1]. In line with the approach proposed in [2], our work aims at finding homogeneous subpopulations of AD patients in which it will be easier to identify statistically and logically relevant relations between clinical and biological descriptors. This is not an easy task because we are looking for homogeneous clusters in a noisy domain with a large set of descriptors with unreliable values of

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

Y. Guo et al. (Eds.): BIH 2015, LNAI 9250, pp. 57–66, 2015.
DOI: 10.1007/978-3-319-23344-4_6

the clinical data and with biological data that may depend on various biological processes, including those not related to AD.

The distinguishing property of our multi-layer clustering methodology is that the constructed clusters are homogeneous in both biological and clinical layers simultaneously [3]. Preliminary results with the same methodology demonstrated that it enables construction of coherent clusters of AD patients with biological properties that are interesting for expert medical evaluation [4]. In this work we continue the research by using a significantly extended dataset, especially with regard to the number of included clinical and biological properties of patients.

The main novelty of the presented experiments is that we use a gender specific approach. The motivation arose from the preliminary experiments with the same dataset that resulted in some clusters characteristic for either male or female subpopulations. Our primary goal is not the detection of gender related differences in respect of incidence rate or severity of AD, but elimination of gender related characteristics that potentially interfere with the properties related to the dementia. It must be noted that it could be useful to eliminate other sources of variability in biological and clinical data in the same way (e.g., age or ethnic group), but this inevitably reduces the size of populations being analysed. We have therefore concentrated on gender specific analysis only.

The rest of the paper is structured as follows. Section 2 describes the data set and the multi-layer clustering methodology. Section 3 presents the results and statistical comparison of two selected clusters constructed for the male population. Finally, in Section 4 we analyse medical relevance of the results.

2 Data and Methodology

All experiments were performed on the data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database[1]. From the ADNIMERGE dataframe we have extracted the baseline evaluation data for 317 female and 342 male patients. The patients are described by 56 biological and 187 clinical properties. Some numerical values were transformed in order to avoid highly nonlinear variables. Biological descriptors include ABETA peptides, TAU and PTAU proteins, the APOE4 related genetic variations, PET imaging results FDG-PET and AV45, MRI volumetric data (Ventricles, Hippocampus, WholeBrain, Entorhinal, Fusiform gyrus, Middle temporal gyrus (MidTemp)), intracerebral volume (ICV), and results of various laboratory measurements like red blood cells and total bilirubin. Clinical descriptors include Alzheimer's Disease Assessment Scale (ADAS13), Mini Mental State Examination (MMSE), Rey Auditory Verbal Learning Test (RAVLT immediate, learning, forgetting, percentage

[1] The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations. The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California, San Francisco. More information can be found at http://www.adni-info.org and http://adni.loni.usc.edu.

of forgetting), Functional Assessment Questionnaire (FAQ), Montreal Cognitive Assessment (MOCA) and Everyday Cognition, which is a cognitive functions questionnaire filled out by patients (ECogPt) and their study partners (ECogSP) (Memory, Language, Visuospatial Abilities, Planning, Organization, Divided Attention and the Total score), Neuropsychiatric Inventory Questionnaire, Modified Hachinski Ischemia Scale, and Geriatric Depression Scale. As an indication of the medical diagnosis we have used global clinical dementia rating score that is interpreted as clinically normal (CN value 0), mild cognitive impairment (MCI value 0.5) and Alzheimer's disease (AD value 1).

2.1 Multi-layer Clustering

Clustering is a standard machine learning approach but it still suffers from problems such as optimal selection of the distance measure and the number of resulting clusters. Typically, the obtained clusters are unstable because they significantly depend on user selectable parameters. This is especially true for noisy domains and domains with statistically related descriptors (attributes). Recently, a novel approach called *multi-layer clustering* has been developed that successfully solves some of these basic problems [3]. In this approach, the quality of the resulting clusters is ensured by the constraint that clusters must be homogeneous simultaneously in two or more data layers, i.e., two or more sets of distinct data descriptors. By defining the clinical descriptors as one data layer and biological descriptors as the second layer, one can expect not only more reliable clusters, but clusters which will be potentially good candidates for the detection of relevant relations between the clinical and biological descriptors.

The multi-layer clustering consists of two steps. In the first step, for each data layer separately, pair-wise similarity of examples is estimated. In the second step these similarity estimations are used in order to construct clusters. The example similarity table (EST) is an $N \times N$ symmetric matrix, where N is the number of examples. One EST is constructed for each layer by defining an artificial binary classification problem, which needs to distinguish between original data examples and some randomly generated ones. These artificial problems are solved with a supervised machine learning algorithm such as decision trees or rules. These types of models enable us to determine if two examples were classified in the same way, i.e., if they fall in the same leaf or are covered by the same rule. An EST is constructed by counting how many times a pair of original examples is classified in the same way, and these counts represent a similarity estimation between the two examples. In the final EST the counts are normalized.

The second step of the algorithm is the agglomerative clustering. It starts with each example being in a separate cluster and then the algorithm iteratively tries to merge clusters together. In each iteration for each possible pair of clusters the potential variability reduction, based on the EST values that can be obtained by merging the clusters, is computed. Let x_{ij} be the similarity between examples i and j from the EST matrix. The CRV score of a single example i from cluster C is the sum of within cluster and outside of cluster components: $\mathrm{CRV}(i) = \mathrm{CRV}_{\mathrm{wc}}(i) + \mathrm{CRV}_{\mathrm{oc}}(i), i \in C$. The two components are sums of squared deviations

Table 1. List and short descriptions of constructed clusters.

Cluster ID	Number of patients	Distribution of CD diagnoses AD / MCI / CN	Clinical status	Biological properties (with z-score versus cognitive normal)
Clusters for female patients				
F1	46	18 / 24 / 4	Significant cognitive problems (high ADAS13, high FAQ, high MMSE)	low FDG 9.25 high AV45 7.66 low Entorhinal 7.58 low MidTemp 7.45 low Fusiform 6.89 high TAU 6.84
F0	18	0 / 4 / 14	Mild or no dementia	Nothing specific
Clusters for male patients				
M1	20	12 / 8 / 0	Significant cognitive problems	low FDG 6.52 low Hippocampus 5.52 low MidTemp 5.26 high TAU 5.24 low Entorhinal 5.08 low Whole brain 4.89
M2	18	13 / 3 / 2	Significant cognitive problems	low FDG 5.70 high AV45 4.95 low Hippocampus 3.96 high ICV 3.76 high Ventricles 3.61 high APOE4 3.57
M0	20	0 / 1 / 19	Mild or no dementia	Nothing specific

from the mean value within (or outside of) cluster C: $\mathrm{CRV}_{\mathrm{wc}}(i) = \sum_{j \in C}(x_{ij} - \overline{x_{i,wc}})^2$ and $\mathrm{CRV}_{\mathrm{oc}}(i) = \sum_{j \notin C}(x_{ij} - \overline{x_{i,oc}})^2$. Finally, the CRV score of cluster C is the mean value of $\mathrm{CRV}(i)$ values of all examples in the cluster: $\mathrm{CRV}(C) = \sum_{i \in C} \mathrm{CRV}(i)/|C|$. In each iteration for each possible pair of clusters we compute the potential variability reduction that can be obtained by merging the clusters. The variability reduction of joining clusters C_1 and C_2 is computed as: $\mathrm{DIFF}(C_1, C_2) = (\mathrm{CRV}(C_1) + \mathrm{CRV}(C_2))/2 - \mathrm{CRV}(C_2 \cup C_2)$. The pair of clusters with the largest variability reduction in all layers is then merged together. The iterative process repeats until no pair of clusters exists for which the variability reduction is positive. More details on the algorithm can be found in [3].

3 Clustering Results

Table 1 presents the clusters constructed with the multi-layer clustering, independently for the populations of 317 females and 342 males. For the female population we have one large cluster F1 in which the majority of patients have significant problems with dementia. Out of the 46 included patients, 18 have the

diagnosis of AD while 24 have been diagnosed with mild cognitive impairment. In the entire dataset there are 22 patients with AD diagnosis and 18 of them are included into this cluster. The clinical properties of these patients are high ADAS13, FAQ, MMSE scores, and all types of cognitive problems. The biological properties of these patients are also typical for AD patients, e.g., low FDG values, high AV45 values, and significantly decreased Entorhinal volume. A statistical comparison with a population of all 145 female patients with cognitive normal status in the dataset has been used to identify the most distinguishing biological properties of the cluster. The last column of Table 1 presents the most significant properties in respect of the highest z-score values of the Mann-Whitney test. The values are very high denoting that differences between cognitive normal patients and those included in the cluster are very significant.[2] It is surprising that, according to the global clinical dementia rating score, besides 18 AD patients, the cluster also includes 24 patients with the mild cognitive impairment and even 4 cognitive normal patients.

The second cluster F0 constructed for the female population includes 18 patients that are typical patients with no or early mild dementia and have no biological properties that are significantly different from cognitive normal patients. Interestingly, the cluster is relatively small compared to the previous female cluster F1, especially if we take into account that there are 145 cognitive normal female patients in the data set. A probable explanation could be that among ADNI patients diagnosed as cognitive normal there are also patients that are not completely healthy, but their subjective or objective problems are either not severe enough or their problems are in discrepancy with typical clinical profiles.

The bottom part of Table 1 presents clusters for the male population. First, we have a cluster of cognitive normal patients (M0) that is very similar in size and properties to the female cluster F0. The most significant difference to the female population is that there are two male clusters of AD patients (M1 and M2) that have clinical properties of typical AD patients. In the first one (M1) there are 20 patients, 12 of them with AD and 8 with mild cognitive impairment. In the second cluster (M2) there are 18 patients, 13 of them with AD diagnosis.

Although that at the first glance the biological properties characterizing patients in clusters M1 and M2, in contrast to the cognitive normal patients, seem similar, there are substantial differences between these two clusters. Biological and clinical properties that most significantly differ between the clusters according to the Mann-Whitney test are listed in Table 2.

4 Analysis of the Results

Cluster M2 deserves special attention due to the fact that the average values of ICV and whole brain volume for patients in M2 are higher than average values for the set of all 124 cognitive normal male patients. The result is surprising because it is in contradiction with common knowledge that atrophy of human brain is related with cognitive problems and in contradiction with results of structural

[2] A value of a z-score higher than 3.29 denotes statistical significance of $P < 0.001$.

Table 2. Biological and clinical properties that are most significantly different for patients in clusters M1 and M2.

Property	Average value for cognitive normal males	Average value for M2	Average value for M1	Mann-Whitney z-score M1 versus M2
Biological properties				
ICV (*1000)	**1576**	**1728**	**1457**	4.14
Whole brain (*1000)	**1109**	**1154**	**961**	3.95
MidTemp	21640	19748	17000	3.38
Fusiform	19600	18237	16175	3.20
TAU	60	86	126	2.53
Red blood cells	62	159	215	2.17
Hippocampus	7792	6565	5684	2.06
Clinical properties				
Abstraction_moca	1.81	1.78	1.15	2.69
MMscore	29.01	24.5	22.35	2.53
Naming_moca	2.92	2.94	2.25	2.38
Q4score (Delayed Word Recall)	3.35	7.83	9.30	2.34
FAQTV	0.10	1.67	3.00	2.19
RAVLT.immediate	44.15	26.28	19.90	2.02

MRI that for AD patients typically show a pattern of decreased grey matter in different regions of the brain [5]. The differences are statistically significant, average ICV values are 1576 and 1728 for cognitive normal and M2 patients, respectively (z-score 3.76, $P < 0.001$), while average whole brain volumes are 1109 and 1154 (z-score 1.99, $P < 0.05$).[3] When comparing patients in cluster M2 with patients in M1, who also have typical AD symptoms but, as expected, decreased ICV and whole brain volumes, then the differences are even more statistically significant (see Table 2).

The importance of the discovery is manifold. First, it indicates gender specific differences because such a cluster with similar properties is not detected in the female population. Second, for a domain in which biological processes with opposite manifestations (decrease and increase of ICV) may result in similar clinical consequences (dementia), segmentation of the patient population is suggested before other analyses aimed at the discovery of relations between biological and clinical properties of patients are performed. Finally, the result is intriguing in respect of its biological and medical interpretation.

It is possible that the increased ICV and whole brain volumes are a consequence of an artefact in data collection procedures, feature extraction from images, or data post-processing (normalization). The assumption may stimulate careful evaluation of the ADNI data, especially for patients in cluster M2. But the result may also suggest the existence of a different biological pathway

[3] Actual absolute values for ICV and whole brain volume are 1000 times larger.

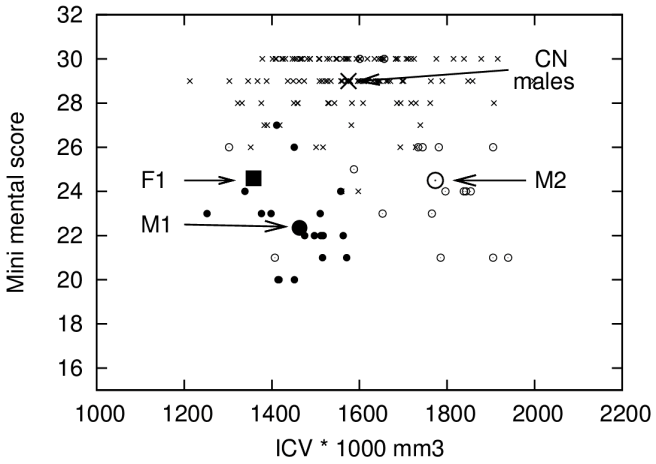

Fig. 1. Patients in cluster M1 (black circles), M2 (white circles), and cognitive normal males (×-marks) presented in respect of intracerebral volume (ICV) and Mini Mental total score values. The large circles and ×-mark denote median values of these clusters and of the cognitive normal male population, respectively. The large black square denotes the mean of female AD patients in cluster F1.

for the male population, resulting in serious dementia problems that are often diagnosed as Alzheimer's disease but with less expressed clinical symptoms (see bottom part of Table 2). In the scientific literature we have found no support for such explanation except that the study devoted to gender related differences [6] concluded that "AD pathology is more likely to be clinically expressed as dementia in women than in men". Fig. 1 illustrates the differences among patients in clusters M1 and M2 and cognitive normal male patients in respect of measured ICV values and Mini Mental scores.

Two male clusters M1 and M2 together include 25 (71%) of a total of 35 male patients with AD diagnosis in the data set. In contrast, the female cluster F1 includes 18 female patients with AD diagnosis (82%) of a total of 22 present in the data set. From the fact that we have one large female cluster and two small

Table 3. Most correlated biological-clinical pairs of properties for various patient populations.

Population	Number of patients	Biological property	Clinical property	Spearman correlation r_s
All	659	FDG	MOCA	–0.51 (df=645, $P < 0.001$)
Female	317	FDG	ADAS13	–0.56 (df=307, $P < 0.001$)
Male	342	Hippocampus	ADAS13	–0.58 (df=289, $P < 0.001$)
F1	46	MidTemp	ADAS13	–0.62 (df=35, $P < 0.001$)
M1	20	TAU	RAVLT.immediate	–0.79 (df=14, $P < 0.01$)
M2	18	ABETA	RAVLT.forgetting	–0.69 (df=9, $P < 0.05$)

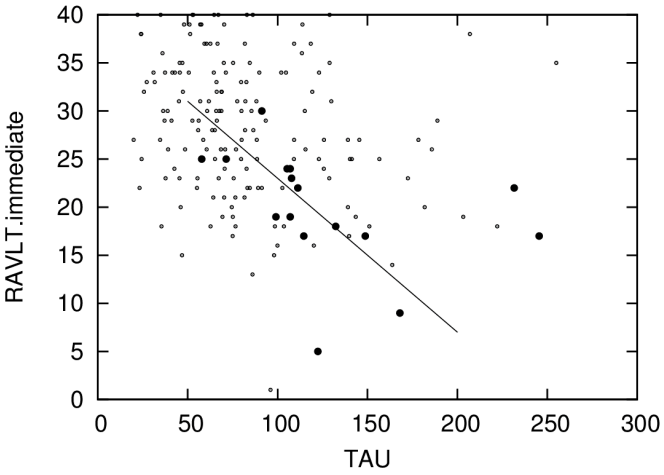

Fig. 2. Patients in cluster M1 (big black circles) and all male patients (small white circles) presented in respect of TAU values and RAVL.immediate scores.

male clusters that together still include a smaller proportion of a population with AD diagnosis than the female one, we can conclude that the male population with serious dementia problems is significantly less homogeneous than the female population. Additionally, the majority of patients in male clusters have AD diagnosis (70% and 80% for M1 and M2, respectively) in contrast to the female cluster F1 in which only 40% of patients have AD diagnosis. The majority of patients in the female cluster have diagnosis of mild cognitive impairment, in spite of the fact that the average values of clinical symptoms are high and biological properties are very different from the cognitive normal female population. This is a surprising finding that is potentially interesting for further analysis, especially in respect of the current diagnostic practice.

One of the stated ADNI goals is to improve clinical trial design through detection of biomarkers that could be used as approximate measures of the severity of dementia. This is known as a difficult task that is still far from a satisfactory solution. If the constructed clusters are really more homogeneous than the complete population, then it may be expected that identification of dementia disease markers should be an easier task for each cluster separately than it is for the complete population. Table 3 presents the best pairs of one biological and one clinical property that can be identified for the complete population, for the female population only, for the male population only, and finally for clusters F1, M1, and M2. The best pairs are identified with the Spearman rank-order correlation coefficient r_s that is computed for all possible pairs of properties. The result confirms that for constructed clusters there exist more strongly correlated biological-clinical relations. The maximal value is detected for cluster M1.[4] The

[4] Only the absolute value is important, the negative sign means an inverse correlation.

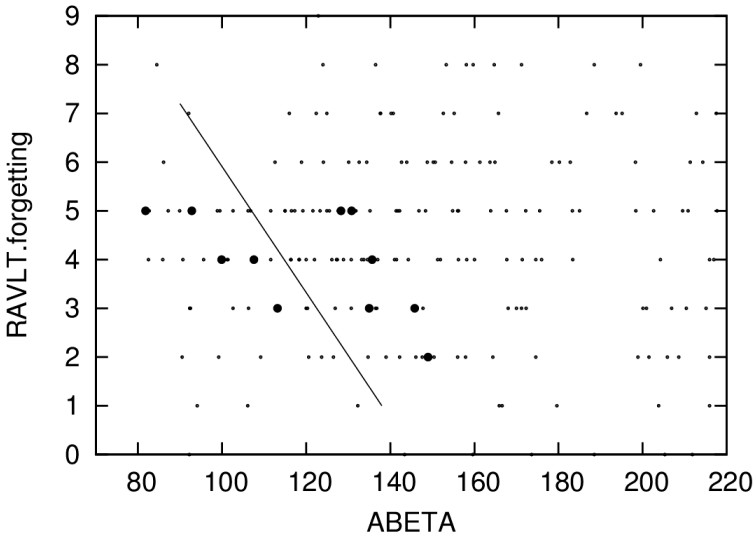

Fig. 3. Patients in cluster M2 (big black circles) and all male patients (small white circles) presented in respect of ABETA values and RAVL.forgetting scores.

correlation is illustrated in Fig. 2 in which patients from cluster M1 and the complete male population are plotted according to their TAU values and RAVLT immediate memory scores. The plotted line is a linear interpolation for patients from cluster M1. In Fig. 3 we present a similar plot of the corresponding best biological-clinical relation for cluster M2. It must be noted that in spite of higher correlation coefficient values, the statistical significance of correlations for small clusters is smaller than for the large cluster because of their size. The result means that detected high correlations are not so reliable and that they have to be confirmed by further experiments. Additionally, it is worth noting the differences between properties that participate in the best pairs. As expected, FDG is the most useful biological property for the general patient population and the result is in agreement with previously reported research [7]. MidTemp, TAU, and ABETA properties are most useful for F1, M1, and M2 clusters, respectively. Also interesting is that RAVLT memory related problems are selected as the most appropriate clinical indicators for both M1 and M2 clusters and that they are correlated with protein specific variables.

5 Conclusions

The presented results confirm that novel machine learning approaches to clustering can indeed be a useful tool for identifying homogeneous patient subsets in various medical knowledge discovery tasks. The applied multi-layer clustering technique and its combination with the gender related separation of the population of patients is definitely not the only possible approach but its results are

promising. Still, significant further research effort in this direction is necessary. Clusters constructed with the multi-layer clustering are small and they contain only a small fraction of all patients. In spite of this, the analysis of the results enabled the conclusion that for Alzheimer's disease there are significant gender specific differences. Additionally, a male subpopulation with a surprising effect of increased ICV and whole brain volume has been detected. The existence of these subpopulations suggests that segmentation of the AD patient population is strongly recommended as a preprocessing step for any analysis aimed at understanding of relations between biological and clinical properties of AD patients, however, based on the available data we still do not know how to practically perform the segmentation in a non ad-hoc manner, especially for the cognitive normal patients and the patients with the mild cognitive impairment diagnosis.

Acknowledgments. We would like to acknowledge the European Commission's support through the Human Brain Project (Gr. no. 604102), MAESTRA project (Gr. no. 612944) and InnoMol project (Gr. no. 316289), support of the Croatian Science Foundation (Pr. no. 9623: Machine Learning Algorithms for Insightful Analysis of Complex Data Structures) and support of the Slovenian Research Agency (program Knowledge Technologies and project Development and Applications of New Semantic Data Mining Methods in Life Sciences). Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from many private pharmaceutical companies and non-profit organizations.

References

1. Weiner, M.W., et al.: The Alzheimer's Disease Neuroimaging Initiative: A review of papers published since its inception. Alzheimer's & Dem. **9**(5), e111–e194 (2013)
2. Galili, T., Mitelpunkt, A., Shachar, N., Marcus-Kalish, M., Benjamini, Y.: Categorize, cluster, and classify: a 3-c strategy for scientific discovery in the medical informatics platform of the human brain project. In: Džeroski, S., Panov, P., Kocev, D., Todorovski, L. (eds.) DS 2014. LNCS, vol. 8777, pp. 73–86. Springer, Heidelberg (2014)
3. Gamberger, D., Mihelčić, M., Lavrač, N.: Multilayer clustering: a discovery experiment on country level trading data. In: Džeroski, S., Panov, P., Kocev, D., Todorovski, L. (eds.) DS 2014. LNCS, vol. 8777, pp. 87–98. Springer, Heidelberg (2014)
4. Gamberger, D., Ženko, B., Mitelpunkt, A., Lavrač, N.: Multilayer clustering: biomarker driven segmentation of Alzheimer's disease patient population. In: Proc. of the International Conference IWBBIO 2015, pp. 134–145 (2015)
5. Hample, H.L., et al.: Biomarkers for Alzheimer's disease: academic, industry, and regulatory perspectives. Nature Reviews Drug Discovery **9**, 560–574 (2010)
6. Barnes, L.L., et al.: Sex differences in the clinical manifestations of Alzheimer disease pathology. Archives of General Psychiatry **62**(6), 685–691 (2005)
7. Langbaum, J.B., et al.: Categorical and correlational analyses of baseline fluorodeoxyglucose positron emission tomography images from the Alzheimer's Disease Neuroimaging Initiative (ADNI). Neuroimage **45**(4), 1107–1116 (2009)