City Research Online

City, University of London Institutional Repository

Citation: Turkay, C., Lundervold, A., Lundervold, A. J. & Hauser, H. (2013). Hypothesis generation by interactive visual exploration of heterogeneous medical data. Paper presented at the Third International Workshop, HCI-KDD 2013, 1 July 2013 - 3 July 2013, Maribor, Slovenia.

This is the unspecified version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/3553/

Link to published version:

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online: http://openaccess.city.ac.uk/ publications@city.ac.uk

Hypothesis Generation by Interactive Visual Exploration of Heterogeneous Medical Data

Cagatay Turkay[1], Arvid Lundervold[2] Astri Johansen Lundervold[3] , and Helwig Hauser[1]

[1] Department of Informatics, University of Bergen, Norway
{Cagatay.Turkay,Helwig.Hauser}@uib.no,
[2] Department of Biomedicine, University of Bergen, Norway
Arvid.Lundervold@biomed.uib.no,
[3] Department of Biological and Medical Psychology, University of Bergen, Norway
Astri.Lundervold@psybp.uib.no

Abstract. High dimensional, heterogeneous datasets are challenging for domain experts to analyze. A very large number of dimensions often pose problems when visual and computational analysis tools are considered. Analysts tend to limit their attention to subsets of the data and lose potential insight in relation to the rest of the data. Generating new hypotheses is becoming problematic due to these limitations. In this paper, we discuss how interactive analysis methods can help analysts to cope with these challenges and aid them in building new hypotheses. Here, we report on the details of an analysis of data recorded in a comprehensive study of cognitive aging. We performed the analysis as a team of visualization researchers and domain experts. We discuss a number of lessons learned related to the usefulness of interactive methods in generating hypotheses.

Keywords: interactive visual analysis, high dimensional medical data

1 Introduction

As in many other domains, experts in medical research are striving to make sense out of data which is collected and computed through several different sources. Along with new imaging methodologies and computational analysis tools, there is a boom in the amount of information that can be produced per sample (usually an individual in the case of medical research). This increasingly often leads to heterogeneous datasets with very large number of dimensions (variables), up to hundreds or even thousands. This already is a challenging situation since most of the common analysis methods, such as regression analysis or support vector machines [1], for example, do not scale well to such a high dimensionality. Consider for instance applying factor analysis to understand the dominant variations within a 500-dimensional dataset. It is a great challenge to correctly interpret the resulting factors even for the most skilled analyst.

On top of this challenge, the number of samples is usually very low in medical research due to a number of factors such as the availability of participants in

a study or high operational costs. This results in datasets with small number of observations (small n) but a very high number of variables (large p). Since most of the statistical methods need sufficiently large number of observations to provide reliable estimates, such "long" data matrices lead to problematic computations [2]. Both the high dimensionality of the datasets and the "$p \gg n$ problem", pose big challenges for the analyst and the computational tools. These challenges lead to the fact that the experts tend to limit their analyses to a subset of the data based on a priori information, e.g., already published related work. Limiting the analysis to a subset of the data dimensions hides relations in the data that can potentially lead to new, unexpected hypotheses.

At this stage, the field of visual analytics can offer solutions to analysts to overcome these limitations [3] [4]. The visual analysis methods enable analysts to quickly build new hypotheses through interaction with the data. The user also gets immediate feedback on whether or not these hypotheses call for a further investigation. Moreover, the interactive tools enable analysts to check for known hypotheses and relationships that have been already studied and reported in the related literature.

In this application paper, we discuss how interactive visual analysis methods facilitate the hypothesis generation process in the context of heterogeneous medical data. We discuss how we utilize the *dual analysis* of items and dimensions [5] in the interactive visual analysis of high dimensional data. We report on the analysis of data related to a longitudinal study of cognitive aging [6] [7]. We demonstrate how our explorative methods lead to findings that are used in the formulation of new research hypotheses in the related study. We additionally showcase observations that are in line with earlier studies in the literature. We then comment on a number of lessons learned as a result of the analysis sessions that we performed as a team of visualization researchers and domain experts.

2 Interactive Visual Analysis Environment

The analysis of the cognitive aging study data is performed through a coordinated multiple view system [8], that primarily makes use of scatterplots. The user is able to make selections in any of the views and combine these selections through Boolean operators, i.e., $\cup, \cap, \neg$. In order to indicate the selections and achieve the focus+context mechanism, we employ a coloring strategy, i.e, the selected points are in a reddish color and the rest is visualized in gray with a low transparency (see Fig. 1-b) to aid the visual prominence of the selection. One additional note here is that we use a density based coloring such that overlapping points lead to a more saturated red color. We use Principal Component Analysis (PCA) – on demand – to reduce the dimensionality of the data when needed. Additionally, we use Multidimensional Scaling (MDS) directly on the dimensions similar to the *VAR display* by Yang et al. [9]. In this visualization approach, the authors represent a single dimension by a glyph that demonstrates the distribution of the items in the dimension. Later authors apply MDS on the dimensions to lay them out on a $2D$-display. Similarly in this work, we feed the correlations

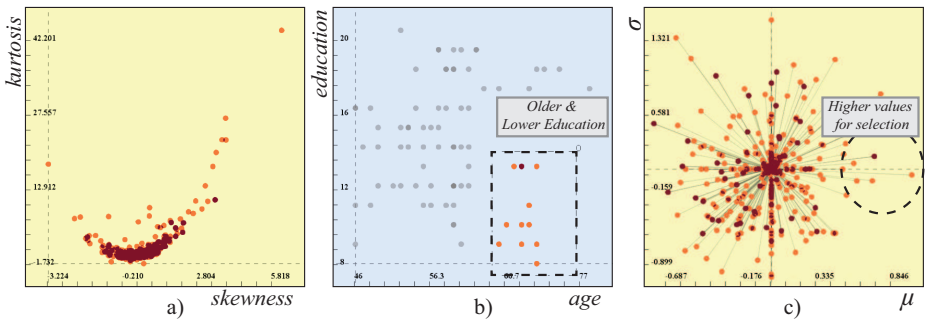

Fig. 1. Dual analysis framework where visualizations of items have a blue and those of dimensions a yellow background. a) We employ a visualization of the dimensions over their *skewness* and *kurtosis* values, where each dot represents a single dimension b) We select a group of participants who are older and have a lower education. c) The deviation plot shows how the μ and σ values change when the selection in (b) is made.

between the dimensions as a distance metric to MDS and as a result, it places the highly inter-correlated groups close to each other. These computational analysis tools are available through the integration of the statistical computation package R [10].

The analysis approach employed in this paper is based on the dual analysis method by Turkay et al. [5]. In this model, the visualization of data items is accompanied by visualizations of dimensions. In order to construct visualizations where dimensions are represented by visual entities, a number of statistics, such as mean (μ), standard deviation (σ), *median*, inter-quartile-range (IQR), *skewness*, and, *kurtosis* are computed for each dimension (i.e., column of the data). These computed statistics are then used as the axes of a visualization of dimensions. In Fig. 1-a, the dimensions are visualized with respect to their skewness and kurtosis, where each dot here represents a dimension.

An additional mechanism we employ is the *deviation plot*, which enables us to see the changes in the statistical computations for dimensions in response to a subset selection of items [11]. In Fig. 1-b, we select a sub-group of participants (from the study data) who are older and have a lower education. We now compute the μ and σ values for each dimension twice, once with using all the items (participants) and once with using only the selected subset. We then show the difference between the two sets of computations in a deviation plot (Fig. 1-c). The dashed circle shows the dimensions that have larger values for the selected subset of items, i.e., for the elderly with lower education. Such a visualization shows the relation between the selection and the dimensions in the data and provides a quick mechanism to check for correlations. Throughout the paper, the views that show items have blue background and those that visualize the dimensions have a yellow background. Further details on the methods could be found in the related references [5] [11].

3 Cognitive Aging Study Data

We analyze the data from a longitudinal study of cognitive aging where the participants were chosen among healthy individuals [6] [7]. All the participants were subject to a neuropsychological examination and to multimodal imaging. One of the expected outcomes of the study is to understand the relations between image-derived features of the brain and cognitive functions in healthy aging [7]. The study involves 3D anatomical magnetic resonance imaging (MRI) of the brain, followed by diffusion tensor imaging (DTI) and resting state functional MRI in the same imaging session [12] [13]. In this paper, we focus on the anatomical MRI recordings together with the results from the neuropsychological examination. The examination included tests related to intellectual function (IQ), memory function, and attention/executive function. IQ was estimated from two sub tests from the Wechsler Abbreviated Scale of Intelligence [14]. The total learning score across the five learning trials of list A (learning), the free short and long delayed recall and the total hits on the Recognition scores from the California Verbal Learning Test (CVLT) II [15] were together with the subtest Coding from Wechsler Adult Intelligence Scale-III [16] used to assess memory function. The Color Word Interference Test from the Delis-Kaplan Executive Function System [17] and the Trail Making Test A and B from the Halstead-Reitan Test Battery [18] were used to assess attention/executive function.

The resulting dataset from the study contains information on 82 healthy individuals who took part in the first wave of the study in 2004/2005. T1-weighted MRI images were segmented into 45 anatomical regions. For each segmented brain region, seven features were derived automatically, namely: *number of voxels*, *volume* and *mean, standard deviation, minimum, maximum and range of the intensity values in the regions*. All these automated computations were done in the FreeSurfer software suite [19]. This automated process creates $45 \times 7 = 315$ dimensions per individual. Additional information on the participants, such as age and sex, and, the results of two neuropsychological tests are added to the data. With this addition, the resulting dataset has 373 dimensions, i.e., the resulting table's size is 82×373. Moreover, meta-data on the dimensions is also incorporated. This meta-data contains whether each dimension is a test score or a brain segment statistic, which brain regions that dimension is related to, and, which statistical feature (e.g., volume or mean intensity) is encoded.

4 Analysis of Cognitive Aging Study Data

In this study, our analysis goal is to determine the relations between age, sex, neuropsychological test scores, and the statistics for the segmented brain regions. The conventional routine to analyze this dataset is to physically limit the analysis to a subset of the dimensions and perform time-consuming, advanced statistical analysis computations on this subset, e.g., loading only the data on specific brain regions and training a neural network with this data. In this setting, if the same analysis needs to be applied on a slightly different subset (which is

often the case), all the operations need to be redone from the beginning – a considerably long time to build/evaluate a single hypothesis. On the contrary, in our interactive methods, the whole data is available throughout the analysis and analysts switch the current focus quickly through interactive brushes.

In order to direct the analysis, we treat age, sex, and the test scores as the dependent variables and try to investigate how they relate to the imaging based variables. Moreover, we investigate the relations within the brain segments. In each sub-analysis, we derive a number of observations purely exploratively. We then discuss these findings as an interdisciplinary team of visualization researchers, experts in neuroinformatics and neuropsychology. We comment on the observations using a priori information and suggest explanations/hypotheses around these new findings. These hypotheses, however, needs to be confirmed/rejected through more robust statistical and/or clinical tests to be considered for further studies. Our aim here is to enable analysts to generate new hypotheses that could potentially lead to significant findings when careful studies are carried out.

Prior to our analysis we handle the missing values and perform normalization on the data. To treat missing values, we apply one of the methods known as statistical imputation and replace the missing values with the mean (or mode) of each column [20]. We continue with a normalization step where different normalization schemes are employed for different data types. Here, dimensions related to the imaging of the brain are z-standardized and the rest of the columns are scaled to the unit interval.

Inter-relations in Test Results. We start our analysis by looking at the relations between the test scores. We first focus our attention on the results related to IQ & Memory function and attention/executive functions related tests and apply a correlation-based-MDS on the 15 dimensions. The rest of the dimensions are not used in the computation and are placed in the middle of the view and colored in gray in Fig. 2-a. Here, we choose to focus on the two large groups, that are to the left and to the right of the view. For a micro analysis, one can focus on the sub-groupings that are visible in both of the clusters. The first group relates to test results assessing IQ and memory function (Group-1). The second group relates to test scores assessing attention and executive function (Group-2). This grouping is in line with the interpretation of these scores and we investigate these two sub-groups separately in the rest of the analysis. We interactively select these sub-groups and locally apply PCA on them. We then use the resulting principal components (PC) to represent these two groups of test scores. We observed that for both of the groups much of the variance is captured by a single PC, so we decide to use only the first PC for each group.
Hypothesis 1: There are two dominant factors within the test results, *IQ & memory* and *attention & executive function.*

Findings Based on Sex. As a continuation of our analysis, we now focus on available meta-data on patients, such as age and sex, to derive interesting

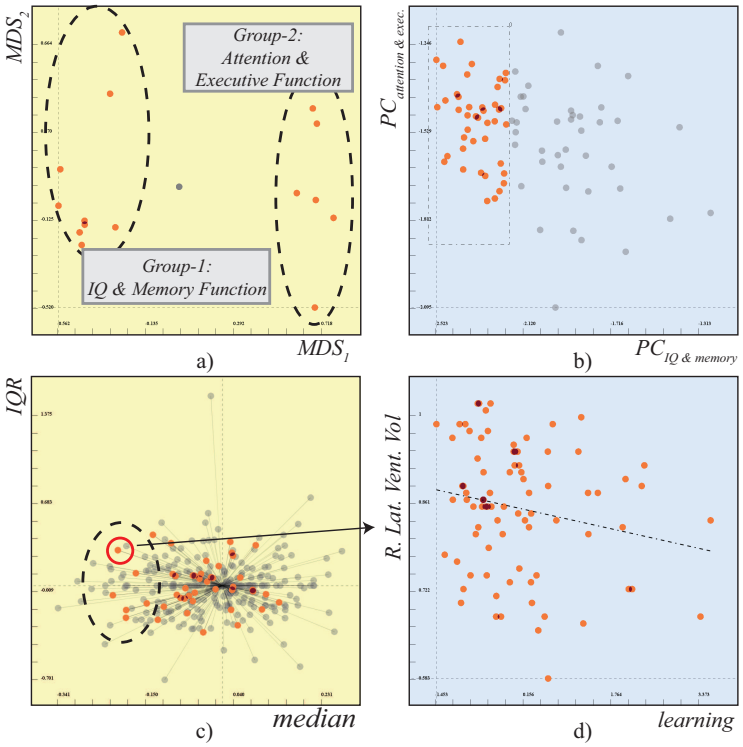

Fig. 2. a) MDS is applied on the *test score* dimensions, where related dimensions are placed close to each other. Two groups for the test scores (Group-1: IQ and memory related, Group-2: attention) show up in the results. b) Each group is represented through an application of PCA and the resulting first principal components are mapped to the axes of the scatterplot. A group of participants, who are better in learning and attentive function is selected. c) Some brain regions are smaller for this subgroup, i.e., have smaller *median* value. d) We select one of the dimensions that shrink the most, *right lateral ventricle volume* (red circle), and visualize these values against the learning scores from CVLT. We notice that there is indeed a negative correlation with the learning score from the CVLT.

relations. We begin by a visualization of *age* vs. *sex* and select the male participants (Fig. 3-a) with a brush and observe how the test scores change in the linked deviation view (Fig. 3-b). The visualization shows that the male participants performed worse in *IQ & memory function* related tasks. In tests related to attention and executive function, however, there were no significant changes between sexes. This is a known finding that has been already observed throughout the study. Another observation that is also confirmed by prior information is the differences in brain volumes between sexes. An immediate reading in Fig. 3-c is that *male participants have larger brains (on average) compared to women*,

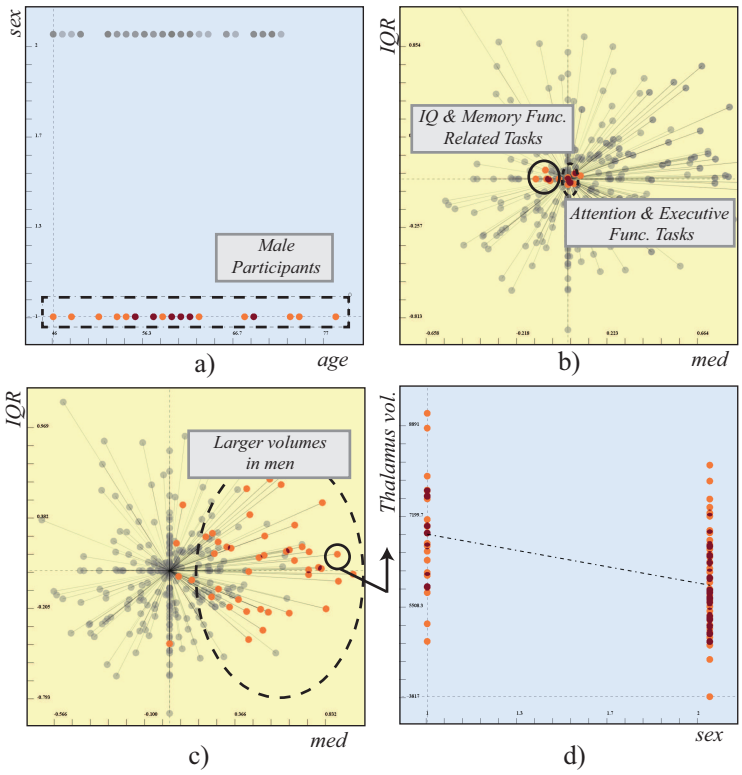

Fig. 3. Male participants are selected (a) and the deviation plot shows that for IQ & memory related tasks, males generally perform worse. However, for attentive and executive function related tests, there is no visible difference (b). When the changes in volume for the brain segments are observed, it is clearly seen that males have larger brains (c). When the volume of one of the segments, thalamus, is visualized with a linear regression line, the sex based difference is found to be significant.

which is a known fact. We analyze further by selecting one of the regions that changed the most, *Thalamus volume*, and look at its relation with sex (Fig. 3-d). We see that there is a significant change, however, this apparent sex difference in thalamic volume has shown to be negligible when the intracranial volume (ICV) difference between sexes are taken into account [21]. This finding could probably be further explored by normalizing segmented brain volumes with the subject's ICV (if this measure is available).

Hypothesis 2: Males perform worse in *IQ & memory* related tests but not in those related to *attention & executive function*.

Findings Based on Age. We continue our investigation by limiting our interest to the elderly patients to understand the effects of aging on the brain and the

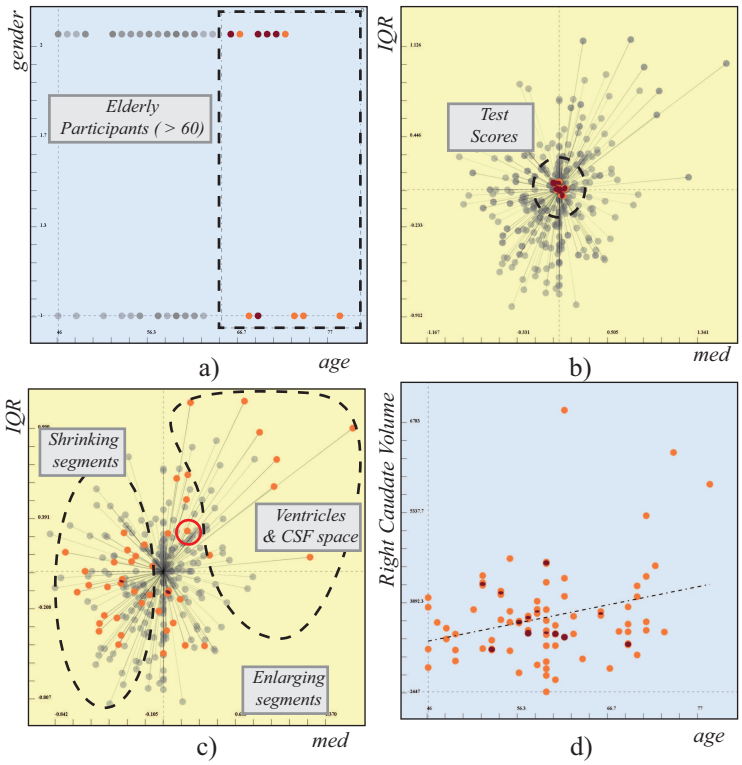

Fig. 4. Elderly patients (> 60 years old) are selected (a). No significant relation is observed in the test scores (b). When we focus on the volumes of the segments, we see most of the regions are shrinking with age, but some, especially the ventricles, are enlarging (c). Apart from the expected enlargement of the ventricles, *the right caudate* is also found to enlarge with age (d).

test results. We select the patients over the age of 60 (Fig. 4-a) and visualize how brain volumes and test scores change. We observed no significant difference in IQ & memory and attentive functions for the elderly patients (Fig. 4-b). However, when we observe the change in brain volumes, we observe that there is an overall *shrinkage in most of the brain segments with age*. This is clearly seen in Fig. 4-c, where most of the dimensions have smaller *median* values (i.e., to the left of the center line). Although most of the brain regions are known to shrink with age [22], some regions are reported to enlarge with age. When the dimensions that have a larger *median* value due to the selection (i.e., enlargement due to aging) are observed, they are found to be the *ventricles* (not the *4th ventricle*) and the *CSF space*. Since this is a known fact [22], we focused on the regions that shows smaller enlargements and decide to look at *the right caudate* more closely. When *the right caudate* is visualized against age, a significant correlation is ob-

served (Fig. 4-d). This is an unexpected finding that needs to be investigated further.
Hypothesis 3: There is no significant relation between age and performance in IQ & memory and attentive & executive functions for individuals undergoing a healthy aging. Moreover, in contrast to the most of the brain regions, there is a significant enlargement in *the right caudate* in healthy aging individuals.

IQ & Memory Function vs. Brain Segment Volumes. We oppose the first principal components for the two groups of test scores (Fig. 2-a) and select the participants that show better IQ & memory function performance (Fig. 2-b). A linked deviation plot shows the change in *median* and IQR values where we observe the change in the imaging related variables (Fig. 2-c). We limit our interest to the variables that are the *volumes of the brain segments* by selecting the volume category through a histogram that displays the related meta-data (not shown in the image). In the deviation plot, we see a sub-group of segments (dashed circle) that have lower volumes for the selected participants (i.e., those that showed better performance). Among those segments are the lateral ventricles that show a significant change. Lateral ventricles are filled with cerebrospinal fluid and have no known function in learning and IQ. We use the integrated linear regression computation on a scatterplot showing *learning* vs. *right lateral ventricle volume* and observe that there is in fact a negative correlation. This could be explained such that, when the ventricles have larger sizes, it indicates less gray matter volume in the brain parenchyma responsible in cognitive function, and is thus associated with reduced performance in IQ & memory function. However, although ventricles tend to grow with age, we observed no significant relation between aging and the performance (See Hypothesis 3). These are now two related observations that leads to an interesting hypothesis.
Hypothesis 4: Regardless of age, the larger sizes of the ventricles are associated with low performance. However, the (expected) enlargement of the ventricles with aging does not directly influence the overall performance.

Relations within Brain Segments. We continue by delimiting the feature set for the brain regions to their *volume* and apply MDS on the 45 dimensions (one for each segment) using the correlation between the dimensions as the distance metric. We identify a group of dimensions that are highly correlated in the MDS plot (Fig. 5-a). This group consists of the volumes for different *ventricles* (lateral, inferior) and *non-white matter hypointensities*. We investigate this finding closely by looking at the relations between *left lateral ventricle* and *non-WM-hypointensities* and found a positive correlation relation (Fig. 5-b) due to a sub-group of patients that have outlying values. This is an interesting finding since non-white matter hypointensities (as segmented by FreeSurfer) might represent local lesions in gray matter such as vascular abnormalities that have a predilection for involving the thalamus and the basal ganglia. Such vascular abnormalities in deeper brain structure could then lead to substance loss and enlarged lateral ventricles. One might further expect that this pathophysiological

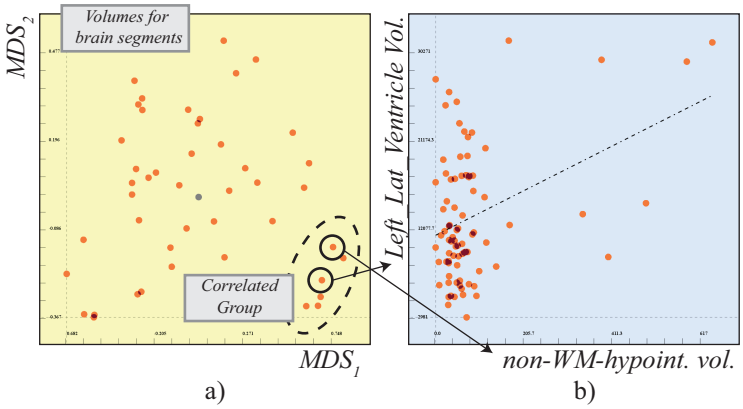

Fig. 5. After MDS is applied on the volume dimensions for brain segments, a correlated group of brain segments is observed (a). Although most of these dimensions are related to the volume of different parts of *the ventricles* (which is expected), *non white matter hypointensities* (scars on the white matter) is also related. This is an interesting finding which led to an hypothesis on the relation between the enlargement of the scars on the white matter and the ventricles.

process would be increasingly frequent with age, but such relationship between age and non-white matter hypointensities was observed to be insignificant in our analysis.

Hypothesis 5: There is a positive relation between lesions on brain tissue and the volume of the ventricles. However, no significant relation with such lesions and age has been detected, this is likely due to the fact that the study involves only participants going through healthy aging.

5 Discussions, Lessons Learned & Conclusions

In a typical analysis of this data, domain experts usually utilize complex machine learning methods, such as neural networks [1], to analyze the data and confirm hypotheses. With such methods however, the process is not transparent and the results can be hard to interpret.

Explorative methods, such as this one presented here, offers new opportunities in building hypotheses. However, the hypotheses built in such systems may suffer from over-fitting to the data, i.e., the finding could be a great fit for a specific selection but harder to generalize [23]. In order to provide feedback on this problem of over-fitting, interactive systems could include cross-validation (or bootstrapping) functionalities to report on the sensibility of the results [24]. In these methods, the hypotheses are tested for several subsets of the data to check the validity of the findings [24]. Another important feature that needs to be present in such interactive systems is the immediate use of more robust and solid

statistical verification methods. In our current framework, we employ linear regression to check for the statistical significance of certain relations (see Fig. 4-d). Such functionalities, and even more advanced inferential statistics, are feasible to incorporate through the embedding of R. Such extensions are desirable for domain experts and can increase the reliability of the results considerably in interactive frameworks.

In this work, we only employed scatterplots and the deviation plot. One can easily extend the selection of visualizations using more advanced methods discussed in the literature. The changes can be encoded by flow-based scatterplots [25] and the comparison of groups can be enhanced by using clustered parallel coordinates [26].

In a significantly short analysis session, we were able to build 5 hypotheses from the healthy aging data. Building this many potential hypotheses using the conventional analysis process would require a considerable amount of time. Throughout the analysis, we discovered relations that lead to novel hypotheses for the healthy aging domain. In addition, we came up with a number of findings that have been already confirmed in the related literature.

6 Acknowledgments

We would like to thank Peter Filzmoser for the valuable insights on the statistical foundations of this work. The study on cognitive aging was supported by grants from the Western Norway Regional Health Authority (# 911397 and #911687 to AJL and 911593 to AL).

References

1. Mitchell, T.M.: Machine learning. McGraw Hill series in computer science. McGraw-Hill (1997)
2. Chen, J., Chen, Z.: Extended bic for small-n-large-p sparse glm. Statistica Sinica **22**(2) (2012) 555
3. Keim, D., Mansmann, F., Schneidewind, J., Thomas, J., Ziegler, H.: Visual analytics: Scope and challenges. Visual Data Mining (2008) 76–90
4. Kehrer, J., Ladstädter, F., Muigg, P., Doleisch, H., Steiner, A., Hauser, H.: Hypothesis generation in climate research with interactive visual data exploration. IEEE Transactions on Visualization and Computer Graphics (IEEE TVCG) **14**(6) (Oct 2008) 1579–1586
5. Turkay, C., Filzmoser, P., Hauser, H.: Brushing dimensions – a dual visual analysis model for high-dimensional data. IEEE Transactions on Visualization and Computer Graphics **17**(12) (dec. 2011) 2591 –2599
6. Andersson, M., Ystad, M., Lundervold, A., Lundervold, A.: Correlations between measures of executive attention and cortical thickness of left posterior middle frontal gyrus - a dichotic listening study. Behavioral and Brain Functions **5**(41) (2009)
7. Ystad, M., Lundervold, A., Wehling, E., Espeseth, T., Rootwelt, H., Westlye, L., Andersson, M., Adolfsdottir, S., Geitung, J., Fjell, A., et al.: Hippocampal volumes

are important predictors for memory function in elderly women. BMC medical imaging **9**(1) (2009) 17
8. Ward, M.O.: Xmdvtool: integrating multiple methods for visualizing multivariate data. In: Proceedings of the conference on Visualization '94. VIS '94, IEEE Computer Society Press (1994) 326–333
9. Yang, J., Hubball, D., Ward, M., Rundensteiner, E., Ribarsky, W.: Value and relation display: Interactive visual exploration of large data sets with hundreds of dimensions. Visualization and Computer Graphics, IEEE Transactions on **13**(3) (may-june 2007) 494 –507
10. Team, R.D.C.: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. (2009)
11. Turkay, C., Parulek, J., Hauser, H.: Dual analysis of dna microarrays. In: Proceedings of the 12th International Conference on Knowledge Management and Knowledge Technologies. i-KNOW '12 (2012) 26:1–26:8
12. Hodneland, E., Ystad, M., Haasz, J., Munthe-Kaas, A., Lundervold, A.: Automated approaches for analysis of multimodal mri acquisitions in a study of cognitive aging. Comput. Methods Prog. Biomed. **106**(3) (June 2012) 328–341
13. Ystad, M., Eichele, T., Lundervold, A.J., Lundervold, A.: Subcortical functional connectivity and verbal episodic memory in healthy elderly – a resting state fmri study. NeuroImage **52**(1) (2010) 379 – 388
14. Wechsler, D.: Wechsler abbreviated scale of intelligence. Psychological Corporation (1999)
15. Delis, D.C.: California verbal learning test. Psychological Corporation (2000)
16. Wechsler, D.: Wechsler adult iintelligence scale-iii (wais-iii). Psychological Corporation, San Antonio (1997)
17. Delis, D.C., Kaplan, E., Kramer, J.H.: Delis-Kaplan executive function system. Psychological Corporation (2001)
18. Reitan, R., Davison, L.: Clinical neuropsychology: current status and applications. Series in Clinical and Community Psychology. Winston (1974)
19. FreeSurfer: http://surfer.nmr.mgh.harvard.edu (2012)
20. Scheffer, J.: Dealing with missing data. Research letters in the information and mathematical sciences **3**(1) (2002) 153–160
21. Sullivan, E.V., Rosenbloom, M., Serventi, K.L., Pfefferbaum, A.: Effects of age and sex on volumes of the thalamus, pons, and cortex. Neurobiology of aging **25**(2) (2004) 185–192
22. Walhovd, K.B., Fjell, A.M., Reinvang, I., Lundervold, A., Dale, A.M., Eilertsen, D.E., Quinn, B.T., Salat, D., Makris, N., Fischl, B.: Effects of age on volumes of cortex, white matter and subcortical structures. Neurobiology of aging **26**(9) (2005) 1261–1270
23. Hawkins, D.M., et al.: The problem of overfitting. Journal of chemical information and computer sciences **44**(1) (2004) 1–12
24. Kohavi, R., et al.: A study of cross-validation and bootstrap for accuracy estimation and model selection. In: International joint Conference on artificial intelligence. Volume 14., Lawrence Erlbaum Associates Ltd (1995) 1137–1145
25. Chan, Y.H., Correa, C.D., Ma, K.L.: Flow-based scatterplots for sensitivity analysis. In: Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on, IEEE (2010) 43–50
26. Johansson, J., Ljung, P., Jern, M., Cooper, M.: Revealing structure within clustered parallel coordinates displays. In: Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on. (2005) 125–132