Fuzzy Clustering to Asses BALI and LIBRA factors for Estimation of DTI measures

Ahmad Akbarifar
Department of Mechanical Engineering, Science and Research Branch, Islamic Azad University
Tehran, Iran
ahmad.akbarifar@srbiau.ac.ir

Adel Maghsoudpour
Department of Mechanical Engineering, Science and Research Branch, Islamic Azad University
Tehran, Iran
a.maghsoudpour@srbiau.ac.ir

Fatemeh Mohammadian
Department of Psychiatry Roozbeh Hospital, Tehran University of Medical Sciences
Tehran, Iran
fmohammadianr@sina.tums.ac.ir

Morteza Mohammadzaheri
Department of Engineering Birmingham City University
Birmingham, UK
morteza.mohammadzaheri@bcu.ac.uk

Omid Ghaemi
Department of radiology and interventional radiology, imam khomeini hospital (imaging center) and shariati hospital Tehran University of Medical Sciences
Tehran, Iran
oghaemi@sina.tums.ac.ir

***Abstract*—Diffusion magnetic resonance imaging (dMRI) is a popular technique for diagnosing dementia through finding a number of measures with diffusion tensor imaging (DTI). However, this technique is too expensive to be widely used to scan populations. The primary objective of this research is to identify factors/indices which are both (i) rather inexpensive to find, and (ii) usable to estimate DTI measures and eventually to diagnose dementia. This will the basis for a low-cost diagnostic solution. Such factors are selected amongst lifestyle for brain health (LIBRA) and brain atrophy and lesion index (BALI) factors. These factors are pertinent to dementia and relatively inexpensive to find. However, BALI and LIBRA are comprised of 49 factors altogether, and development of a diagnostic algorithm with 49 inputs is infeasible. Therefore, it is necessary to pick the most impactful factors to be used in diagnosis algorithm development. Fuzzy subtractive clustering was employed for this purpose. This research shows that the grey matter lesions and subcortical dilated perivascular spaces (GM-SV) and periventricular white matter lesions (PV) from BALI and age, level of education, job status, antidepressant drugs, diabetes control drugs, obesity (BMI) and dementia preventive diet from LIBRA are the most influential factors to identify DTI measures.**

Keywords— ***Fuzzy subtractive clustering, Dementia, LIBRA, BALI, DTI, Diffusion MRI***

I. INTRODUCTION

Dementia, including Alzheimer's disease, is a serious risk for independence and quality of life, particularly for the ever-increasing aged population, i.e. people 65+ years old [1]. With the current trend, globally, the number of dementia patients is estimated to reach 66 million by 2030 and 115 million by 2050 [2].Nowadays, dementia is a substantial emotional, physical, and financial burden on patients, carers, health systems and societies[3]. Dementia will soon cause serious socio-economic and health-system issues for all countries [4, 5]. According to a 2018 report, Alzheimer, the most prevalent form of dementia, imposes a cost of over $350,000 a year in the United States, as the most costly medical condition in the country [6]. Comparably, in Sweden's healthcare system, the cost of dementia is at the same level of cancer and heart disease [7]. In addition to direct medical expenses for dementia patients, e.g. consultations with medical experts, specialist visits, and hospitalizations, indirect costs, such as the lowered productivity of both dementia patients and their carers should be considered too [8].

There is a growing interest in developing early detection methods that can identify dementia before it leads to cognitive deficits [8]. The significance of early detection methods in preventing the onset of dementia is well recognized, and this has led to a demand for robust and reliable methods that can accurately identify the disease in its early stages [8]. Therefore, it is imperative to explore various diagnostic techniques, including neuroimaging and symptom interpretation, to improve the accuracy of diagnosis and to enhance the ability to control the disease [8]. Widely accepted diagnosis technique for dementia is based on the information gained from dMRI and processed through DTI [9]. However, as a serious challenge, unfortunately, dMRI is too facility-demanding and expensive for the entire population. Costly diagnosis based on DTI makes some patients give up the diagnosis process [10]. In addition, the expense of the diagnosis through DTI can put substantial financial pressure on public health systems [10].

Therefore, screening of a sizable fraction of the population needs an alternative low-cost approach. The advent of inexpensive means for identifying individuals who may potentially suffer from cognitive impairment and offering them extensive examinations would be a long step forward in the fight against dementia [11, 12]. Such low-cost cognitive status evaluations would be especially beneficial in underdeveloped countries, whereby vital resources can be conserved for high-risk patients [13, 14]. Gradual progression of dementia. provides an opportunity for early diagnosis while the patient's cognitive abilities are still manageable, and the disease can be halted through treatment.

Patient's lifestyle factors (LIBRA) as well as BALI factors, extracted from a simple MRI, can be seen as relatively inexpensive to attain indicators of disease onset [13, 14]. However, as shown in Tables 1 and 2 in Appendices, total number of BALI and LIBRA factors that could potentially serve as early indicators of dementia is 49. Interpretation of such a vast number of factors and their potential effect on dementia is impractical. Human expert cannot accurately analyse such a high quantity of factors. The real solution is the development of computerized data-driven

model/algorithm to use these factors as its inputs. However, a model which receive all 49 BALI and LIBRA factors would be so large, with numerous parameters, and need medical records for thousands of patients to be developed. Therefore, to pave the way for the development of reliable and universal models/algorithms for the early detection of dementia, it is imperative to identify the most useful BALI and LIBRA factors for dementia diagnosis as a primary step.

II. Problem Statement

This research, with use of collected clinical data, aims to identify a number of BALI and LIBRA factors which are most usable to estimate DTI measures. The reason is, as explained in the introduction, (i) DTI measures, which are expensive to acquire, are widely accepted to be straightforwardly used in dementia diagnosis I [15], and BALI and LIBRA factors, which are numerous but inexpensive to attain, are known to be relevant to dementia. These factors are comprised of 42 life style brain health (LIBRA) risk factors [15, 16], listed in Appendix 1 and 7 brain atrophy and lesion indices (BALI) [17], listed in Appendix 2.

In this research, fuzzy subtractive clustering was employed to select salient BALI and LIBRA factors to estimate DTI measures. This technique identifies the factors that display the most drastic changes across all subjects (patients) as the factors of choice. Conversely, the factors that exhibit little variation amongst subjects are recommended to be excluded from inputs list of any model or algorithm to estimate DTI measures. In the literature, Analysis of Variance (also known as ANOVA) have been widely reported to carry out such tasks [18]. In this research, due to some intrinsic issues of ANOVA [19], fuzzy subtractive clustering was selected to be tried. Prior to use of fuzzy subtractive clustering, the collected data should be prepared.

III. Data Collection

The data were collected from a cohort of 51 elderly subjects from diverse regions of Iran. These subjects were referred to the Memory Clinic of Roozbeh Hospital in Tehran, and were chosen out of the 201 volunteers who were eligible to participate. It is worth noting that several subjects were excluded from the study due to their (1) inability to undergo the tests, (2) absence of interest in continuing the tests, and (3) the practitioner's ultimate decision to remove them from the study. The collection of LIBRA data, relevant to demographic and clinical characteristics, lifestyle factors, and medication usage, were collected through face-to-face interviews by professionally trained researchers. As a part of data collection protocol, investigators were carefully trained to ensure the utmost accuracy and consistency. To assess factors relating to brain health, four categories of LIBRA lifestyle factors were examined, which included: (1) socio-economic demographic factors, (2) factors associated with physical health, (3) lifestyle factors, and (4) laboratory factors, all of which are detailed in Appendix 1. Moreover, BALI factors, listed in Appendix 2, denote morphological changes, and their data were gathered with MR imaging [20].

IV. Data Pre-processing

After examinations and imaging of the subjects, numerical representation of each LIBRA or BALI factor, as outlined in Appendices 1 and 2, was determined for each subject. Nevertheless, the values of these factors exhibit distinct ranges, which in turn, which is drawback and may prevent an impartial analysis [21]. To circumvent this issue, all the data of LIBRA and BALI factors as well as DTI measures were standardized with (1), so that their range changed to [0 1].

$$x_s = \frac{x\text{-}\min(X)}{\max(X)\text{-}\min(X)} \,. \tag{1}$$

where x is the numeric value assigned to each factor, X represents the set of all x values, and s denotes the standardized value.

V. Fuzzy Clustering

A linear Sugeno-type fuzzy inference system (FIS), a class of fuzzy systems, that have been proven to be effective in data-driven modelling, was considered as the potential model in this research [22, 23]. Neuro-fuzzy networks can be built from these fuzzy systems [24].

Let us assume that the FIS has n rules and m inputs; the model's output is a weighted sum of the outputs of all the rules, where each rule receive all m inputs ($m \leqslant 49$ in this research). Each input in every rule is associated with a membership function, which calculates membership grade or μ. In this research, the membership function associated with i^{th} input in the j^{th} rule, is a Gaussian function:

$$\mu_{ij} = \exp\left(-\frac{(u_i - C_{ij})^2}{2\Omega_{ij}{}^2}\right). \tag{2}$$

where u_i is the i^{th} input to the model and C_{ij} and Ω_{ij} denote the centre and the width (standard deviation) of the membership function associated with i^{th} input and the j^{th} rule. The weight of each rule is determined by the product of its membership grades, which range between 0 and 1 and appear in the denominator of equation (3). Furthermore, in linear Sugeno fuzzy systems, the output of each rule is a linear combination of its inputs, as shown in the numerator of equation (3).

$$\textit{model output} = \frac{\sum_{j=1}^{n}\left(\overbrace{\left(\sum_{i=1}^{m} A_{ij}u_i + B_j\right)}^{j^{th}\ \textit{rule output}}\prod_{i=1}^{m}\mu_{ij}\right)}{\sum_{j=1}^{n}\underbrace{\prod_{i=1}^{m}\mu_{ij}}_{j^{th}\ \textit{rule weight}}}. \tag{3}$$

Prior to identification of the model, i.e. finding elements of A, B, C, and Ω, the number of inputs and rules, m and n respectively, should be identified, and this task is the focus of this paper.

As a challenges, if all the inputs had an identical number of membership functions, the number of fuzzy rules would equal n^m rules [25]. As an instance, in this research, if all 49 BALI and LIBRA factors are used as input to the fuzzy model and only three membership functions are assigned to each input, the model would have 3^{49} or nearly 2.39×10^{23} rules. Knowing that each rule would have 49 membership

function and 98 parameters, i.e. C_{ij} and Ω_{ij} s. The model would have 2.35×10^{23} parameters only in membership functions. In order to identify these many parameters, billions of datasets would be required, which are evidently unavailable. The realistic solution is to identify less influential inputs through subtractive clustering and remove them from the list.

Subtractive clustering begins with calculating the density value of each set of input data or $\mathbf{u}^k$ from pre-processed data; k refers to a subject or patient, and $\mathbf{u}^k$ includes 49 BALI and LIBRA factors:

$$D_k = \sum_{l=1}^{51} \exp\left(-\frac{\left\|\mathbf{u}^k - \mathbf{u}^l\right\|^2}{\left(r_a/2\right)^2} \right), \tag{4}$$

where,

$$\left\|\mathbf{u}^k - \mathbf{u}^l\right\| = \sqrt{\sum_{i=1}^{49} (\mathbf{u}_i^k - \mathbf{u}_i^l)^2}\ , \tag{5}$$

and r_a = Range of Influence (a positive number).

The point with the highest density is designated as the centre of the first cluster and the data point with the density smaller than "Reject Ratio" are eliminated from the process. Then, the density of other points is redefined:

$$D_k = D_k - D_{C1} \exp\left(-\frac{\left\|\mathbf{u}^k - \mathbf{u}^{C1}\right\|^2}{\left(r_a/2\right)^2} \right), \tag{6}$$

where r_b is called Squash Factor. If redefined density of any data point falls behind "Reject Ratio", it is eliminated, and if it exceeded "Accept Ratio", it is defined as the centre of a new cluster; and the density of other points is redefined again:

$$D_k = D_k - \sum_{t=1}^{\substack{number\ of \\ exisiting\ clusters}} D_{Ct} \exp\left(-\frac{\left\|\mathbf{u}^k - \mathbf{u}^{Ct}\right\|^2}{\left(r_a/2\right)^2} \right), \tag{7}$$

This process continues until the clusters remains unchanged between two consecutive stages. Clusters may be considered as spheres with the diameter of r_a in a 49-dimension space. Each of these clusters builds membership functions of a fuzzy rule. The centre of j^{th} cluster in the dimension of i^{th} input will be C_{ij} in (2), while $\Omega_{ij}=r_a$.

The pseudocode below describes the process of developing the FIS:

- Find the density of each data point, $\mathbf{u}^k$ with (4 and 5), where $k \in \mathbb{N} \cap [1\ 51]$, 51 is the number of subjects (patients).

10 - Choose the data point with highest density, in which its density is higher than "Accept Ratio" as the centre of a new cluster.

- Number of clusters after 10 is assigned to NC
- *if* NC=ONC then *go to* 20
- NC is assigned to ONC or old NC.
- Eliminate the data points (subjects) with the density<Reject Ratio from the process.
- Subtract the influence of the existing cluster with (7), which is a generalised form of (6).
- *go to* 10

20 – Make a blank Sugeno FIS with same number of rules as the number of clusters.

- A BALI or a LIBRA factor in the centre of each cluster is the centre of a Gaussian membership, (2), where its width equals r_a in (4 and 5). Thus, 49 membership functions are made for each cluster to be used in a fuzzy rule associated to the cluster.

In fuzzy clustering, each rule has a membership function per input (a BALI or LIBRA factor), or each input has the same number of membership function as the number of fuzzy rules or clusters. However, some membership functions of an input are so close to each other that are almost identical. If such membership functions are considered as one, each input ends up with a number of distinct membership functions. Higher number of distinct membership functions for an input is a sign that clusters may cover a wider range in that input, and that input is more influential in estimating DTIs (or any other output). Theoretically, if an input has only one membership function, it means that in the centres of all data clusters, that input has the same value. Such an input plays the same role in all fuzzy rules and can be ignored in modelling; however, it does not necessary mean it is not an important factor in the disease. An input which consistently influence the outputs (i.e. DTI measures) in the same consistent way may be actually very important in the disease but not much useful to estimate DTI measures (or other outputs, which may be considered in the future).

VI. Result and Analysis

Subtractive clustering with Aspect Ratio of 0.5, Reject Ratio of 0.15, Squash Factor of 1.25 and Range of Influence of 0.5 led to 40 clusters/fuzzy rules. Membership functions, in which the distance of their centres is greater than 0.02, were considered as distinct membership functions. The number of distinct membership function for each BALI or LIBRA factor was found and depicted in Figs. 1 and 2. Following factors have at least 10 distinct membership functions: age, level of education, job status, antidepressant drugs, diabetes control drugs, obesity (BMI) and dementia preventive diet from LIBRA factors as well as grey matter lesions and subcortical dilated perivascular spaces (GM-SV) and periventricular white matter lesions(PV) from BALI. These 9 factors are good candidates to be the inputs to a model to estimate DTI measures for Iranian population.

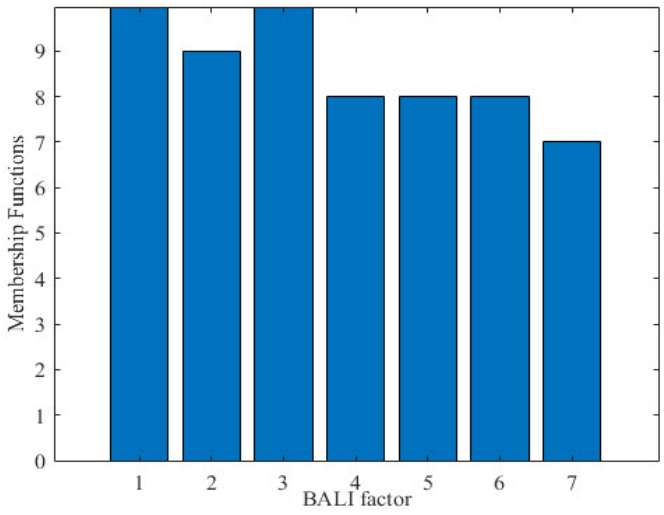

Figure 1. Membership Functions Diagram of Fuzzy Subtractive Clustering Analysis for BALI factors

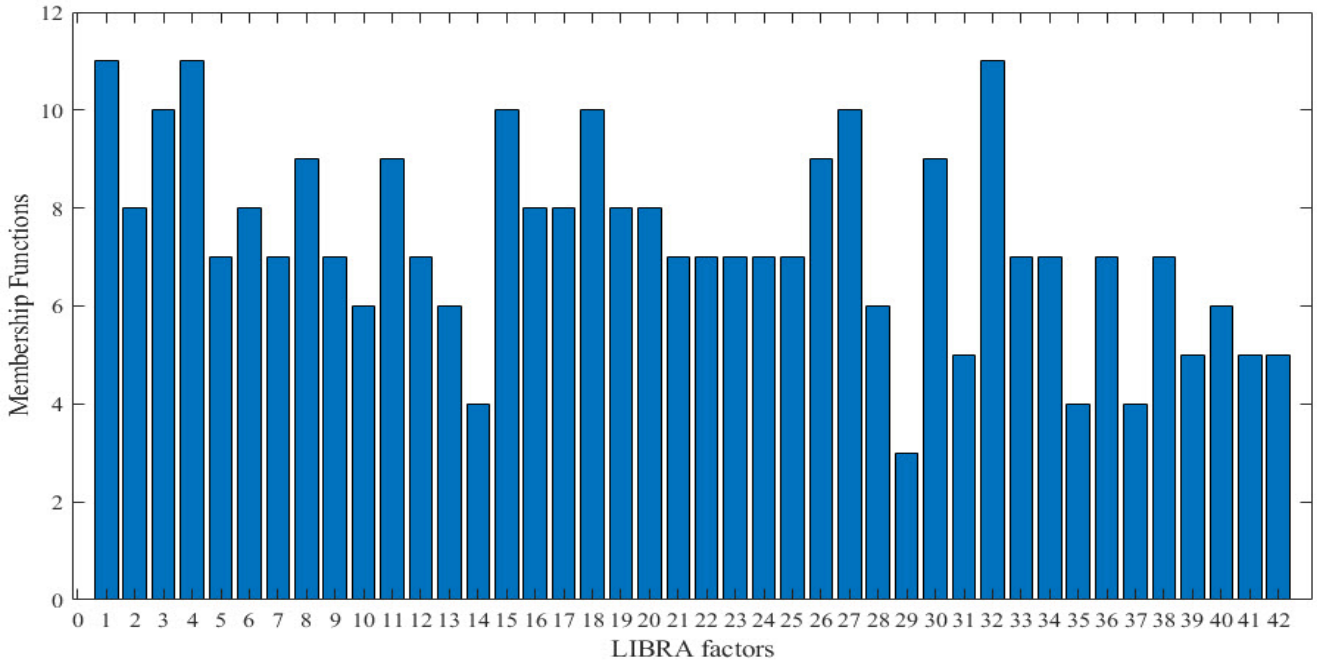

Figure 2. Membership Functions Diagram of Fuzzy Subtractive Clustering Analysis for LIBRA factors

On the other hand, the factors with the smallest number of distinct membership functions exhibit a low variation amongst subjects and can be considered fairly consistent for dementia patients. The factor with the fewest membership function is "saturated fat intake" which can be considered as a serious risk factor for Dementia.

VII. CONCLUSION

An inexpensive diagnostic method is a necessity to avoid significant negative effects of this disease. Currently, DTI measures, calculated based on dMRI results, have proved to be accurate and effective. However, DTI is too expensive to be used for large populations. An alternative is to find a number of relevant inexpensive-to-find factors and mapping those factors to DTI measures with a model/algorithm. Such a model could eventually estimate DTI measures, or practically diagnose dementia with use of aforementioned inexpensive-to-find factors. This research suggests the LIBRA factors as those inexpensive-to-find factors.

However, BALI the LIBRA include 49 factors overall, and this research shows use of all of these inputs leads to having a gigantic model with numerous parameters. Finding parameters of such a model requires access to the records of a too large population. The alternative is to pick a number of factors with highest influence to estimate DTI measures (and theoretically any other relevant output). Fuzzy subtractive clustering was employed for this purpose, which practically assigned same number of fuzzy membership functions to each factors, in which some of those membership functions are too close and were assumed to be identical. The factors with the highest number of distinct membership functions were picked as the candidates to be used in development of a data-driven model to estimate DTI measures. These factors show wider variation amongst the tested patients, volunteered from all over Iran. It should be noted that the chosen factors are not necessarily the most effective factors in development of the disease.

ACKNOWLEDGMENT

The authors wish to especially thank at Advanced diagnostic and Interventional radiology research centre (ADIR), Imam Khomeini complex hospital،Tehran university of medical science, Tehran, Iran, for their kind support of this project.

REFERENCES

[1] D. G. Blazer, K. Yaffe, and J. Karlawish, "Cognitive aging: a report from the Institute of Medicine," *Jama,* vol. 313, no. 21, pp. 2121-2122, 2015.

[2] W. H. Organization, *Dementia: a public health priority*. World Health Organization, 2012.

[3] Z. Feng, L. Coots, Y. Kaganova, and J. Wiener, "Hospital and emergency department use by people with Alzheimer's disease and related disorders," *Hospital,* vol. 8, p. 01, 2013.

[4] B. Reisberg, M. B. Shulman, C. Torossian, L. Leng, and W. Zhu, "Outcome over seven years of healthy adults with and without subjective cognitive impairment," *Alzheimer's & Dementia,* vol. 6, no. 1, pp. 11-24, 2010.

[5] M. Rönnlund, A. Sundström, R. Adolfsson, and L.-G. Nilsson, "Subjective memory impairment in older adults predicts future dementia independent of baseline memory performance: Evidence from the Betula prospective cohort study," *Alzheimer's & Dementia,* vol. 11, no. 11, pp. 1385-1392, 2015.

[6] D. Mukherji, M. Mukherji, N. Mukherji, and A. s. D. N. Initiative, "Early Detection of Alzheimer's Disease with Low-Cost Neuropsychological Tests: A Novel Predict-Diagnose Approach Using Recurrent Neural Networks," *medRxiv,* p. 2021.01. 17.21249822, 2021.

[7] A. Miley-Akerstedt *et al.*, "Lifestyle factors are important contributors to subjective memory complaints among patients without objective memory impairment or positive neurochemical biomarkers for Alzheimer's disease," *Dementia and geriatric cognitive disorders extra,* vol. 8, no. 3, pp. 439-452, 2018.

[8] P. Colligris, M. J. Perez de Lara, B. Colligris, and J. Pintor, "Ocular manifestations of Alzheimer's and other neurodegenerative diseases: the prospect of the eye as a tool for the early diagnosis of Alzheimer's disease," *Journal of ophthalmology,* vol. 2018, 2018.

[9] A. Marcisz and J. Polanska, "Can T1-Weighted Magnetic Resonance Imaging Significantly Improve Mini-Mental State Examination-Based Distinguishing Between Mild Cognitive

Impairment and Early-Stage Alzheimer's Disease?," *Journal of Alzheimer's Disease,* no. Preprint, pp. 1-17, 2023.
[10] E. Choi, M. T. Bahadori, A. Schuetz, W. F. Stewart, and J. Sun, "Doctor ai: Predicting clinical events via recurrent neural networks," in *Machine learning for healthcare conference*, 2016, pp. 301-318: PMLR.
[11] V. Berti, C. Polito, G. Lombardi, C. Ferrari, S. Sorbi, and A. Pupi, "Rethinking on the concept of biomarkers in preclinical Alzheimer's disease," *Neurological Sciences,* vol. 37, pp. 663-672, 2016.
[12] B. Dubois *et al.*, "Preclinical Alzheimer's disease: definition, natural history, and diagnostic criteria," *Alzheimer's & Dementia,* vol. 12, no. 3, pp. 292-323, 2016.
[13] P. S. Aisen *et al.*, "On the path to 2025: understanding the Alzheimer's disease continuum," *Alzheimer's research & therapy,* vol. 9, pp. 1-10, 2017.
[14] A. Porsteinsson, R. Isaacson, S. Knox, M. Sabbagh, and I. Rubino, "Diagnosis of early Alzheimer's disease: clinical practice in 2021," *The journal of prevention of Alzheimer's disease,* vol. 8, pp. 371-386, 2021.
[15] I. S. Heger *et al.*, "Associations of the Lifestyle For Brain Health index with structural brain changes and cognition: results from the maastricht study," *Neurology,* vol. 97, no. 13, pp. e1300-e1312, 2021.
[16] I. Heger *et al.*, "Associations of the LIfestyle for BRAin Health (LIBRA) index with structural brain changes and cognition: results from The Maastricht Study," *Towards primary prevention of dementia,* p. 27, 2022.
[17] L. A. Grajauskas, H. Guo, R. C. D'Arcy, and X. Song, "Toward MRI-based whole-brain health assessment: the Brain Atrophy and Lesion Index (BALI)," *Aging Medicine,* vol. 1, no. 1, pp. 55-63, 2018.
[18] J. N. Rouder, M. Schnuerch, J. M. Haaf, R. D. J. C. B. Morey, and Behavior, "Principles of model specification in ANOVA designs," vol. 6, no. 1, pp. 50-63, 2023.
[19] P. I. Good and C. E. J. J. o. M. A. S. M. Lunneborg, "Limitations of the analysis of variance," vol. 5, no. 1, p. 5, 2006.
[20] T. Gu, H. Guo, and X. Song, "Reliability of the MRI-based Brain Atrophy and Lesion Index in the evaluation of whole-brain structural health," *Aging Medicine,* vol. 1, no. 2, pp. 125-132, 2018.
[21] D. A. Freedman, *Statistical models: theory and practice.* cambridge university press, 2009.
[22] M. Mohammadzaheri *et al.*, "Fuzzy Analysis of Resonance Frequencies for Structural Inspection of an Engine Cylinder Block," *Fuzzy Information and Engineering,* vol. 13, no. 3, pp. 266-276, 2022.
[23] M. Mohammadzaheri, A. AlQallaf, M. Ghodsi, and H. Ziaiefar, "Development of a Fuzzy Model to Estimate the Head of Gaseous Petroleum Fluids Driven by Electrical Submersible Pumps," *Fuzzy Information and Engineering,* vol. 10, no. 1, pp. 99-106, 2018.
[24] A. Ghaffari, A. R. Mehrabian, and M. Mohammad-Zaheri, "Identification and control of power plant de-superheater using soft computing techniques," *Engineering Applications of Artificial Intelligence,* vol. 20, no. 2, pp. 273-287, 2007.
[25] M. Mohammadzaheri, S. Grainger, and M. Bazghaleh, "Fuzzy Modeling of a Piezoelectric Actuator," *International Journal of Precission Engineering and Manufacturing,* vol. 13, no. 5, pp. 663-670, 2012.

APPENDIX 1

TABLE 1 The Impact of Lifestyle on Brain Health (LIBRA) [15, 16]

	Factors	**Rating scheme**
	1- Socio-economic demographic components	
1	Age	(50-60=0)(60-70=+1)(70-80=+2)(90-80=+3)(>90=+4)
2	Sex	(F=+2)(M=0)
3	Level of education	(illiterate = +3) (5 years = +2) (12-5 years = +1) (16-12 years = 0)
4	Job	(Unemployed=+3)(Stressful=+2)(Office=+1)(Environment=0)
5	Marital status	(divorced = +2) (single = +1) (married = 0)
6	Income	(<5=+2)(5<10=+1)(10<=0) in million Tomans
	2- Components of physical health	
7	History of diabetes mellitus	(<5year=+1)(5-10year=+2)(>10year=+3)(NO=0)
8	History of blood pressure	(<5year=+1)(5-10year=+2)(>10year=+3)(NO=0)
9	History of high cholesterol	(<5year=+1)(5-10year=+2)(>10year=+3)(NO=0)
10	Kidney dysfunction	(OK=+1)(NO=0)
11	Coronary heart disease	(CABG=+3)(stent=+2)(OK=+1)(NO=0)
12	Benzodiazepine	(<5year=+1)(5-10year=+2)(>10year=+3)(NO=0)
13	Antipsychotic drugs	(<5year=+1)(5-10year=+2)(>10year=+3)(NO=0)
14	Anticholinergic drugs	(<5year=+1)(5-10year=+2)(>10year=+3)(NO=0)
15	Antidepressant drugs	(<5year=+1)(5-10year=+2)(>10year=+3)(NO=0)
16	Antiplatelet drugs	(<5year=+1)(5-10year=+2)(>10year=+3)(NO=0)
17	Fat profile modulating drugs	(<5year=+1)(5-10year=+2)(>10year=+3)(NO=0)
18	Diabetes control drugs	(<5year=+1)(5-10year=+2)(>10year=+3)(NO=0)
19	Cardiovascular drugs	(<5year=+1)(5-10year=+2)(>10year=+3)(NO=0)
20	Family history	(OK=+1)(OK AND DEAD=+2)(<60=+3)(NO=0)
21	History of stroke	(OK=+1)(NO=0)(BIG=+2)(OK=+1)(NO=0)
22	History of psychiatric illness (depression-anxiety-psychosis)	(OK=+1)(NO=0)
23	Interrupted breathing in sleep (sleep apnea)	(PSG=+2)(OK=+1)(NO=0)
24	Hearing health	(OK=+1)(OK AND DON'T USE=+2)(NO=0)
25	Vision health	(OK=+1)(OK AND DON'T USE=+2)(NO=0)
	3- Lifestyle components	
26	Physical activity	(>2H=-1)(1H-2H=0)(30min-1H=+1)(<30min=+2)(NO=+3)
27	Obesity - BMI	(20-25=0)(<20=+1)(25-30=+1)(>30=+2)
28	Smoking	(<30PACK/YEAR=+1)(>30PACK/YEAR=+2)
29	Saturated fat intake	(OLIVE OIL-SESAME=-1)(LIQUID OIL =0)(SOLID OIL=+1)
30	High cognitive activity	(NO=+2)(<1H=+1)(2H-1H=0)(>2H=-1)
31	Alcohol intake	(OK=+1)(NO=0)
32	Dementia preventive diet	(<FISH-vegetables-fruit per 3/week=+1/+1/+1)
33	Chronic lack of sleep	(OK=+1)(NO=0)
34	Having a lot of stress	(OK=+1)(NO=0)
	4-Laboratory components	

35	Examination of C-reactive protein (CRP)	(+=+1)
36	Cobalamin (B12)	(<180=+2)(320-180=+1)(>320=0)
37	Folate/folic acid	(<5=+1)(>5=0)
38	High cholesterol	(>150=+1)
39	Homocysteine	(>20=+2)
40	Fasting blood sugar (FBS)	(>110=+1)
41	HbA1C	(>6.5=+1)
42	TSH	(5<=+1)(USE OF DRUG=+2)

APPENDIX 2

TABLE 2 Neuroimaging Ratings for brain atrophy and lesion index (BALI) [17]

	Factors	**Description**	**Rating schema**
1	***GM-SV***	Gray matter lesions and subcortical dilated perivascular spaces	0=absence; 1=dotted abnormal SI in GM or multiple dotted/liner abnormal SI in subcortical areas; 2=small patches of abnormal SI in GM or diffuse and countless dotted/liner abnormal SI in subcortical areas; 3 =patches of abnormal SI in GM
2	***DWM***	Deep white matter lesions	0=absence; 1=dotted abnormal SI; 2=small patches of abnormal SI; 3=large patchy abnormal SI lesions; 4=large patchy abnormal SI involving all cerebral lobes; 5 =abnormal SI involving complete deep WM
3	***PV***	Periventricular white matter lesions	0=absence; 1="cap"or pencil-thin lining;2=smooth"halo"with blurred margin; 3=irregular periventricular abnormal signal intensities extending into the deep WM
4	***BG***	Lesions in the basal ganglia and surrounding areas	0=absence; 1=1 focal lesion; 2=more than 1 focal lesion; 3=patchy confluent lesions(regardless of dilated perivascular spaces)
5	***IT***	Lesions in the infratentorial regions	0=no obvious atrophy; 1=mild atrophy; 2=moderate atrophy; 3=severe atrophy; 4=most severe atrophy present especially in the medial temporal lobes; 5=most severe atrophy present especially in the medial temporal lobes and cerebral cortex
6	***GA***	Global atrophy	0=absence; 1=1 focal lesion; 2=more than 1 focal lesion; 3=diffuse lesions
7	***MH***	Microhemorrhage	0=no other kind of finding; 1=any one kind; 2=any two kinds; 3=more than two kinds