

The Growing Role of Complex Sensor Systems and Algorithmic Pattern Recognition for Vascular Dementia Onset

Janna Madden¹, Arshia Khan²
Department of Computer Science
University of Minnesota, Duluth
Duluth, USA

Abstract—Vascular Dementia is often Clinically diagnosed once the effects of the disease are prevalent in a person’s daily living routines. However, previous research has shown various behavioral and physiological changes linked to the development of Vascular Dementia, with these changes beginning to present earlier than clinical diagnosis is currently possible. In this review, works focused on these early signs of Vascular Dementia are highlighted. However, recognizing these changes is difficult. Many computational systems have been proposed for the evaluation these early signs of Vascular Dementia. The chosen works have largely focused on utilizing sensors systems or algorithmic evaluation can be incorporated into a person’s environment to measure behavioral, and psychological metrics. This raw data can then be computationally analyzed to draw conclusions about the patterns of change surrounding the onset of Vascular Dementia. This compilation of works presents current a framework for investigating the various behavioral and physiological metrics as well as potential avenues for further investigating of sensor system and algorithmic design with the goal of enabling earlier Vascular Dementia detection.

Keywords—Vascular dementia; pattern recognition; machine learning; artificial intelligence; algorithmic disease detection; vascular dementia onset

I. INTRODUCTION

Vascular Dementia has been considered a distinct diagnosis since 1991, when a consensus was reached for the diagnosis criteria for vascular dementia at the National Institutes of Health [1]. However, the intermingling of vascular events and dementia was recognized far earlier. As knowledge of vascular pathology and behavior increased through the 19th and 20th centuries, researchers came across the association of vascular events and cerebral changes, at the time thought synonymous to the onset of traditional dementia [1, 2]. This finding became a major hurdle in the field-- the quickly advancing research into the realm of vascular disease brought forth the need for research in the field of cerebrovascular disease which proved more elusive [2, 3].

While findings slowed during this time, effort did not. While much research focused on dementia symptoms as a whole, some researchers discussed the effects of vascular events on the progression of dementia, leading to the eventual differentiated vascular dementia from other forms of dementia. Otto Binswanger did just in his 1910 publication that is

considered the origin of the modern study of vascular dementia. In this publication, Binswanger suggested that cerebral impairment was a result of vascular insufficiency [2, 3].

Despite this early proposition, it wasn’t until 1991 that Vascular Dementia was clinically differentiated, and 1995 that Vascular Cognitive Impairment, the precursor of Vascular Dementia, was considered a unique diagnosis [1]. Currently, Vascular Dementia and Vascular Cognitive Impairment are considered diagnosis distinct from other forms of dementia and cognitive impairment respectively. From this, it follows that Vascular Dementia and Vascular Cognitive Impairment have symptoms and onset patterns that are specific to the disease.

Pattern recognition is a realm of computation that has proved very successful in various environments [11]. Given that the onset of Vascular dementia has been associated with particular changes in biometrics, it is proposed that a combination of these biometrics may lead to more precise and accurate diagnosis.

A. Motivation

When it comes to pattern recognition, algorithms and learning systems have proven successful solutions in a broad range of industries including cyber-security, fault detection in safety-critical systems, textual anomaly detection, image analysis (such as facial recognition, object identification or analysis of medical imaging) and biometric detection [11, 12, 13]. One specific method of pattern recognition used for cyber-security is the artificial immune system-- the system seeks to identify and classify changes that happen to itself. Some changes may be considered safe changes while others pose a risk to the system [12]. This is widely applied to cyber-security, however, the evolutionary nature of evaluating the “risk” of a detected occurrence could be applied disease diagnosis as well. In the case of biometric detection methodologies, we can see the distinctive and identifiable nature of biometrics such as gait. It follows that a significant enough change to a biometric measure over time would also be observable [13].

These examples depict how current applications of computational pattern recognition can translate to the problem of disease detection. With that being said, applying pattern recognition strategies to the problem of diagnosing vascular

dementia, and generally healthcare data, involve many of the identified “challenges data features” discussed in the survey of Anomaly Detection by Chandola et al. [11] that contribute to the difficulty of the detection of the pattern recognition problem. Specifically, healthcare data is such that:

- 1) “Normal (Healthy) Region” of specific data measures can vary person-to-person and can even fluctuate in response to other factors.
- 2) Health is an evolutionary domain, in which natural changes such as aging make it so a current representation of “Normal Region” may not be sufficient representation in the future.
- 3) The severity of a fluctuation depends on the metric being used to measure and the meaning associated with that metric. (1 degree of fluctuation in heart rate may be considered ok, whereas the same fluctuation in temperature would be of concern.
- 4) Often data contains join which tends to be similar to anomalies and hence make it difficult to distinguish and remove.
- 5) The cost of incorrectly classifying an anomaly as normal can be very high.

Diagnosing Vascular Dementia using pattern recognition therefore poses many challenges. With that being said, given that Vascular Dementia is first recognizable through biometrics and lifestyle observations, specifically executive functioning tasks, non-invasive sensor systems and pattern recognition tools that are based on daily functioning and biometric data have the potential to identify early traits of dementia onset.

B. Contribution

In this paper, we present a taxonomy for the categorization of current computational pattern recognition developments within the realm of Vascular Dementia onset and progression. Following this, each category is discussed in detail, looking at the innovations and contributions to the particular niche, all within the context of the current Vascular Dementia knowledge base. In included contributions have a focus non-invasive, non-imaging measures that have proposed correlations to vascular dementia onset.

For each category, the unique assumptions that apply in terms of pattern recognition and anomaly detection are discussed and the supposed complexity of the techniques in a wide-scale implementation. We conclude the review with a comparative discussion of all current sub-domains of contributions and future works.

The remainder of this paper is organized as follows: Section 2 briefly reviews vascular dementia progression and diagnosis, Section 3 discusses the major sub-domains of computational pattern recognition, Section 4 presents the taxonomy of pattern recognition techniques in the domain of vascular dementia, Section 5 through 8 discuss the application of pattern recognition in the following sub-domains: cognition and executive functioning assessments; behavioral and emotional assessments; eye tracking; gait analysis; motor control; linguistic patterns; sleep patterns; and health-record mining respectively in relation to supporting early

identification of dementia onset. A discussion and concluding remarks follow.

II. BACKGROUND

A. Vascular Cognitive Impairment and Vascular Dementia

Since Vascular Cognitive Impairment and Vascular Dementia originate from a vascular event, the risk factors are the same as that for cardiovascular disease, including hypertension, stroke, atrial fibrillation, aortic fibrillation, diabetes mellitus type 2, obesity, lack of active lifestyle, depression, sleep apnea and smoking [4]. Within the diagnosis of Vascular Cognitive Impairment and Vascular Impairment, the disease progression can future sub-divided as: Pre-clinical Vascular Cognitive Impairment, Vascular Cognitive Impairment, and Mild, Moderate, Moderately Severe, or Severe Vascular Dementia.

Pre-clinical Vascular Cognitive Impairment: During this initial stage, the cerebral changes have no measurable symptoms displayed-- changes are not detectable on clinical assessments and symptoms are either not noticed or of such a weak intensity that they are diagnosable. Because of this, much of what is known about the pre-clinical stage is learned from retrospective study of diagnosed cases. One study found that patients had memory complaints 12 years prior to diagnosis and had experienced declines in activities of daily living 5 to 7 years previous to diagnosis [5]. While vascular dementia patients had memory complaints 12 years prior to diagnosis, comparable to that of other forms of dementia, there is comparatively less deterioration in the preclinical stage as compared to other forms of Dementia. “Executive Functioning” is a term that encompasses many of the cognitive tasks that experience deterioration throughout onset and as early as the pre-clinical stage including: cognitive flexibility (problem solving through different means, thinking about a situation in different ways), working memory (comprehending and recalling information), and inhibitory control (self-control, ignore distractions, regulated emotions and impulses) [6]. Additionally, mental health concerns arise during the pre-clinical stage including depression, lack of interest or motivation, and loss of energy. This association still remains significant after adjusting for memory complaints, showing that depression symptoms are more than a by-product of perceived cognitive difficulties.

Vascular Cognitive Impairment: The transition from pre-clinical to vascular cognitive impairment is difficult to unanimously identify. Vascular Cognitive Impairment is loosely defined as case where one or more cognitive domains become significantly affected [7, 8]. At this stage of progression, symptoms are sometimes clinically diagnosable and while symptoms can be noticeable in daily living, they are not to limiting in this respect.

Vascular Dementia: Onset of vascular dementia is marked by cognitive impairment severe enough to interfere with everyday activities [9]. The Diagnosis of Vascular Dementia can be subdivided further into mild, moderate, moderately severe, and severe.

Just as onset revolves around a variety of symptoms, there is no single test that can diagnosis Vascular Cognitive

Impairment or Vascular Dementia. Diagnosis is based firstly based on the presence of cognitive impairment and the presence and assumption that cerebrovascular disease is the cause of the present cognitive impairment. If this is the case, assessments of thinking (measured using neuropsychological tests), behavior, daily functioning, neurological reflexes and coordination, brain imaging, and carotid ultrasound are all possible tools used to verify the presence and identify the assumed cause of the symptoms [9, 10]. In addition, medical records play an important role by providing a history record of any past memory complaints and risk factors (as previously discussed) [10]. Identifying the onset of Vascular Dementia is challenging for multiple reasons. Because it is a delayed onset disease, symptoms present in increasing severity, identifying the beginning of onset presents one challenge. Additionally, the disease can present with varying symptoms, resulting in case-by-case variability. Despite the difficulties, early diagnosis is vastly important as it can lead to treatments that reduced the impact and progress of the disease.

B. Developments in Computational Pattern Recognition

Pattern recognition is central to most tasks we perform on a daily basis. Indeed, even at a young age, children can recognize letters and numbers, despite differences in penmanship or fonts [14]. This example reinforces the findings of Herbert Simon, economist by training who's research largely focuses on factors and motivators of decision making and the corresponding outcomes: the greater the number of relevant patterns, the better the resulting decision will be [15, 16]. While on the surface, simple, the idea of discovering and utilizing many relevant patterns is at the heart of most pattern recognition problems. Indeed, in the healthcare realm, and more specifically in the study of delayed chronic disease onset, there is no one stand-alone indicator for disease. Therefore, searching for multiple indicators of disease is necessary so together, this information can be used to formulate a more accurate prediction.

The process of pattern recognition can be described in the following three stages:

- 1) Observing the environment
- 2) Distinguishing patterns of interest from their background
- 3) Concluding reasonable decisions about the categories of the patterns.

In the first step, the agent (either human or computer) uses the data available to it or data that it can aggregate through sensing the environment or making assumptions about the environment to build a conceptual view of the its surroundings. It's important to note that the created view will never completely capture the actual environment [17]. In the process of building a view of the model, the accuracy of each metric and dependency placed on each metric must also be considered. In terms of a human building a view of their environment, an individual might place more trust in one sense, such as sight or hearing, and less determinants in others, depending on the individual's confidence of accuracy in the "data" coming from that particular source. Comparability, in computational models, this weighting is accomplished through

sensor fusion. Sensor fusion methods aim to reduce uncertainty but combining information from various sources in a way that reduces noise variance with the end goal of making the resulting dataset more accurate than the individual data sources used individually [18]. This process of reducing uncertainty in data ties into the following step; distinguishing patterns of interest from their background. As alluded to earlier, the complexity of the environment leads to the necessity of simplification. The second step of pattern recognition is distinguishing meaningful trends from background information. This means evaluating the weight of trust that should be placed on the result of each pattern recognized in the data, eliminating noise and recognizing background data, reducing noise and like mentioned in regard to sensor fusion, reducing uncertainty in data [19]. In the final step, these trends and patterns are used to make the best-available decision. While this process is intuitive and largely automatic for humans, computationally mimicking the pattern recognition process has proven to have individual challenges but, in specific cases, has led to improved outcomes [19, 20]. Various techniques of computational pattern recognition have been theorized or implemented. These techniques can fairly neatly fall into three main groups: intelligent algorithms, statistical models and machine learning.

Intelligent algorithms follow a structured methodology, making decisions based on rules that have been developed based on past data trends. While complex, each decision is traceable through the decision algorithm [21]. Positively, intelligent algorithms have high approval with practitioners because understanding the method of decision making is approachable for most, regardless of background training [22]. With that being said, intelligent algorithms do not scale to growing types of data without intervention. If a greater variety of data is incorporated into the record, the intelligent algorithm would need alterations to utilize the growing knowledge available. Intelligent algorithms are in essence a rule-based system, taking the format: "if a condition holds, do the following". Because of this condition-checking structure of these algorithm, to effectively implement an intelligent algorithm underlying knowledge of the data must exist [21]. Typically, these systems are used to codify known correlations between data values and the outcome in question. Since intelligent algorithms are based on known knowledge, so while they play a role in diagnosis, the discover of new correlations between data is unlikely through this method.

Statistical models build on this idea by describing relationships between variables in mathematical equations and utilize relationships between variables to predict outcomes [23]. Similar to intelligent algorithms, the outcomes of statistical models can be traced to the equations that derived the particular outcome. In this way, statistical models too are approachable across disciplines. In terms of drawbacks, statistical modeling methods have limitations very similar to that of intelligent algorithms: incorporating growing data sources into the model requires human intervention and an underlying knowledge of data correlations is often a prerequisite for building statistical models [23].

Machine learning, a sub-discipline of artificial intelligence and opposed to both intelligent algorithms and statistical

models, machine learning is an algorithm that can learn from data without the use of predefined rules or programming. An additional benefit is that machine learning methods can adapt to different sets of data and growth of data types with minimal alterations [24]. In the realm of medical research where types of data are continually increasing, the ability to grow the model with the growing data source is a notable benefit. Additionally, such models are knowledge of the underlying correlations is not required to build an effective model. Likely, correlations between variables may be suspected, yet not concretely defined. Machine learning models can identify commingling factors and can even lead to knowledge discovery. Machine learning methods come in many varieties, from regression models, Bayesian networks, neural networks and nearest-neighbor based techniques among many, all used to build relationships between variables and predicted outcomes [24, 25]. Machine learning models often suffer in understandability, making the inner-workings of models not readily accessible to outside of the discipline. Despite this drawback, machine learning has shown much potential in the realm of disease diagnosis.

III. TAXONOMY FOR CLASSIFYING TECHNIQUES FOR PATTERN RECOGNITION OF VASCULAR DEMENTIA ONSET

There are many instances of pattern recognition being applied to vascular dementia. Fig. 1 shows our proposed taxonomy for classifying the proposed non-invasive techniques for pattern recognition of Vascular Dementia onset. Current research in the field is categorized by the source of the data. The first of these categories is validated survey-based assessment tools and encapsulates all pattern recognition that is based on data from clinically validated assessments. The second category looks at multiple tools that measure and classify based on biometrics such as gait, balance, motor impairments and eye focus. This realm of research has become especially entwined with technology, as tools are developed to collect biometric data precisely, more frequently and in a more accessible format for patients. Third, tools based on psychological and behavioral based measures. Two primary

areas of research in this subdomain include measuring the difficulty and variety of linguistic choices made by the patient and tracking sleeping patterns. Finally, we look at health record data mining applications as a possible source for data. This category incorporates text mining free-text fields of medical records as well as other temporal analyses of health events recorded in electronic medical records. Concluding remarks discusses the overall progress, limitations and possible future direction of pattern recognition for earlier Vascular Dementia onset detection.

IV. COGNITION AND EXECUTIVE FUNCTIONING ASSESSMENTS

A. Origins of Cognitive and Executive Functioning Assessments

Changes noted in cognition and executive functioning following a vascular event first prompted research into the realm of Vascular Dementia [26, 27]. Current research continues to look at how particular forms of cognition and executive functioning changes indicate the onset of vascular dementia and has led to the creation and use of cognitive assessment tools to measure cognition and executive impairment. Many test such as the WAIS Logical Memory - Delayed Recall Test and Silhouette Naming Tests have been proposed to assess the patient's functioning. In addition to the aforementioned assessments, other assessments tools such as, the Wechsler Memory Scale and Wechsler Memory Scale-Revised, Montreal Cognitive Assessment, Mini-Mental Status Exam, Clock-Drawing Test, Functional Activities Questionnaire, California Learning Test and Hopkins Verbal Learning Assessment have been proposed for the evaluation of cognitive functioning, and applied to vascular dementia research [29, 30, 31]. The assessments vary in their exact methods, length to complete, complexity of administering and availability of parallel forms (which enables retesting without participants recognizing lists from previous testing) -- factors that come into play when researchers decide which assessment to select.

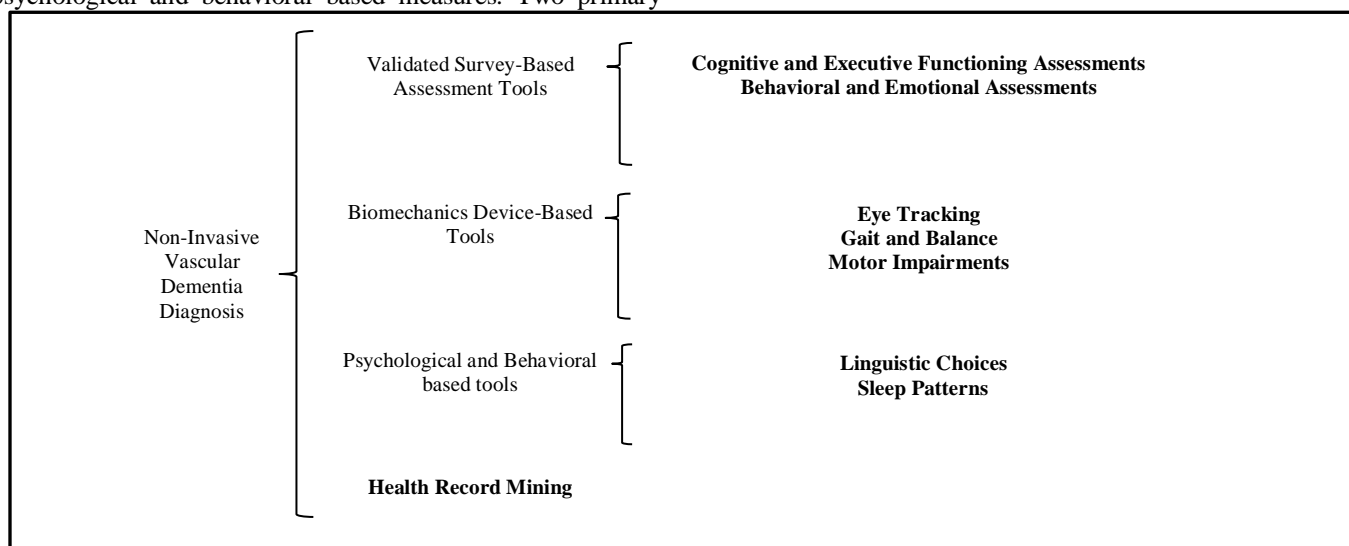


Fig. 1. Topology of Non-Invasive Assessments for Earlier Vascular Dementia Diagnosis. Each Category is Discussed in Detail in the Corresponding Sections Below.

B. Distinguishing forms of Dementia

A study by Grahams, Emery and Hodges shows that using the WAIS Logical Memory-Delayed Recall Test and Silhouette Naming Test cognitive assessments and a logistic regression model on the resulting data, cognitive impairment from vascular dementia was distinguished from cognitive impairment caused by other forms of dementia or cognitive impairments. The findings showed that patients with vascular dementia were more likely to exhibit impairments in semantic memory, executive/attentional functioning and visuospatial and perceptual skills [28].

C. Determine time of Dementia Onset

As demonstrated by the Grahams, Emery and Hodges study [28], these assessments can be used a source for data for building statistical or machine learning models that can help us understand and describe the onset of vascular dementia. Such models can be aimed at distinguishing Vascular Dementia onset from other forms of dementia or distinguishing Vascular Dementia from non-demented patients. One technique for differentiating the populations from Cognitive assessments is error analysis. As noted, Vascular dementia has known characteristic memory impairments. Error analysis looks at the variation in type and ordering of errors in assessment, for example if a test is contains words within three categories, error analysis could look at whether the participant lists uses the categories to recall the words [32, 33].

D. Online Screening for Dementia

Additionally, some online tools have been proposed for the detection of dementia. However, these are limited to confirming the disease after it has been well established. Such tests often revolve around the same principles as cognitive and executive functioning, including verbal recognition, speed of response (reaction time), complex and sustained patterns, and visual memory, only administered in an online format. A unique case is "CogScreen", a functioning assessment created by the Federal Aviation Agency for pilots that has been suggested for patient populations as a way to evaluate early signs of mild dementia. This assessment stands out in that it is made with a healthy population in mind. For the diagnosis of dementia, the goal of identifying the decline from healthy to patient populations may be most easily done using a test designed for a healthy population [34]. While such methods so far have been limited in their specificity of the diagnosis -- therefore the type and severity would not be detectable [34]. In addition, a varied comfort-level with the online-environment may lead to inaccurate conclusions from such tests. However, authors do note the advantages of utilizing including the convenience and low cost for administering.

V. BEHAVIORAL AND EMOTIONAL ASSESSMENTS

Along with cognition, behavioral and emotional changes were also correlated early on with the presence of vascular dementia. As early as 1896, the beginning of decline of vascular dementia was correlated with depressive states as well as psychological agitation, increasing in correlation with the degeneration [28]. The mood of patients was also early linked with regression of vascular dementia, though, even at this stage, it was noted that mood was difficult to evaluate and was

subjective to the reviewer and environmental factors [35]. Indeed, we see that while both behavioral and emotional factors and those of cognitive and executive functioning were being first recognized in regard to what we now know as Vascular Dementia around the same time period, advances in the later has been much more substantial. Behavioral and emotional factors are found to be related to so many diseases tend to be used in conjunction with other diagnosis methodology, though less telling individually. Such factors include increased stress, anxiety, distress and depressive symptoms and reduced apathy [36, 37]. With that being said, modeling tools can be used to provide a tool by which to understand how such behavioral and emotional present in known cases of Vascular Dementia.

VI. BIOMECHANICS: EYE TRACKING

A. Gaze and Visual Search Preferences as an indicator of Cognitive Impairments

Early on in the onset of Dementia. When talking about attention in terms of visual gaze and search, the duration of visual fixation, the number of fixations and the reaction is time to a new stimuli. In a comparative study, participants with Vascular Dementia took longer visually searching for stimuli, especially in visuospatial challenging tasks such as when the stimuli was at a further distance from them or the search field is wider. Such visual impairments characteristic of vascular dementia are primarily rooted in psychomotor slowing, not a specific deficit to the search methods [38]. In terms of analysis, eye-tracking data, algorithms to detect fixations, saccades (rapid switching between two fixations) and backtracking (moving "backwards" to previous fixations) are of particular interest because this enumerates the participation's attention and thus best measure the underlying cognitive impairments associated with neurodegenerative disorders that may be present [39].

B. Eye Tracking as a Measure of Cognitive Functioning

In a visual paired comparison task, eye tracking is used to determine participants' attention choices. When showed two pictures side-by-side, one being seen 2 seconds ago and one being novel, all participants' attention was on the novel one for more than 70% of the time. However, when 2 minutes had elapsed since seeing the image, those with Vascular Cognitive Impairment spend on average 53% on the novel image (and the remainder on the previously seen image) while other participant groups, including a control and a participants with parkinson's disease remained at above 70% of the time on the novel image [40]. Such findings suggests that visual cues can also be used to evaluate memory functioning.

C. Emotional Stimuli Reaction

Emotional processing is also proposed as a key feature for the diagnosis of Vascular Dementia. Participant's scores on emotional recognition tasks have previously correlated with the severity of dementia. Eye tracking can in such tasks to measure the participants attention to particular stimuli. For example, with the choice between negatively, neutrally or positively charged images, participants with vascular dementia were seen to have decreased emotional curiosity as disease onset progressed [41].

VII. BIOMECHANICS: GAIT ANALYSIS

A. Gait Feature Dysfunction

Gait features including gait steadiness, gait speed, balance throughout gait cycle, step frequency, length of single support lines, variability of single and of double support lines, double support time and unsteady gait rhythm have been identified for their possible relation with the onset of Vascular Dementia. In a longitudinal study, these factors were analyzed over time using an in-shoe sensor network to analyze the participant's movements [42, 43]. In terms of predictive modelling, gait analysis often utilizes computational tools to measure features of gait, with predictions and classifications commonly being carried out using statistical models.

B. Dual Tasking on Gait

Aging plays a role in gait analysis. For this reason, experimental design has sought a model that does not revolve simply around gait, which experiences changes associated with many factors of aging. The effect of "distractor" tasks on gait has been examined as a possible means to evaluate a participant's ability to manage both cognitive functioning and motor control simultaneously. During the onset of Vascular dementia, such experiments have shown a decreased ability to dually perform on both tasks [44]. Again, in this area of Vascular Dementia diagnosis, we see limited use of machine learning, with the focus largely on statistical models for a means of drawing conclusions. Authors noted that dual-task disturbances could be indicators of an increased fall risk during early stages of onset.

VIII. BIOMECHANICS: MOTOR CONTROL

Similar to gait analysis, motor control seeks to understand through the physical biomechanics of the body, the functioning of the brain. Motor control more generally than gait analysis, looks at any the control of any muscle movement. Balance, among other aspects of physical functioning can be affected by the onset of Vascular Dementia and can be studied in isolation from gait. A frequently referenced test in Vascular Dementia Diagnosis research is the "finger tapping" test. The goal of such a test is to see the processing or response time that a participant needs to execute a particular task [45]. Some research also looks at daily tasks such as driving; looking at biometrics that are easy to quantify and measure over time, integrates automated collection of data into daily living. Driving assessments significantly separated Vascular Dementia participants from older control group and diabetic control group on driving scores. In addition, driving skills were correlated with performance on the Mini-Mental State Examination [46]. Research in this area has been statistically correlated with vascular dementia, however, like behavioral assessments, the predictive power of these measures is limited by their relation to other aspects of aging.

IX. BEHAVIORAL: LINGUISTIC PATTERNS

A. Linguistic Patterns in Responses to Cognitive Functioning Prompts

Linguistic pattern modeling comes in two general forms: the first is based on the assessment techniques for cognitive functioning, with verbal analysis incorporated, while the

second looks at interpreting linguistic patterns in an unprompted environment. In the case of the first form, vocal markers from assessment prompts are extracted and assessed on the power to distinguish between control, mild cognitive impairment and fully presenting dementia cases [47]. In Konig et al. the Mini-Mental State Examination (MMSE), Five Word Test, Frontal Assessment Battery and Instrumental Activities of Daily Living Scale are used as the neuropsychiatric inventories. Tasks were recorded and vocal features extracted. Factors in the recordings that were considered include: vocal reaction time, relative length to speak a particular sentence (compared to mean of both control and clinical groups), amount of silence, amount of insertions, amount of deletions and irregularities [47].

B. Non-Prompted Linguistic Pattern Recognition

The other primary form of Linguistic Patterns looks at quantifying linguistic choices in a less non-assessment environments. Such tools look to understand spoken text and associate meaning with particular words. This can be as simple as having a "positive" and a "negative" list of various words, but often involves using a language corpus to understand contexts and associate meaning with phrases, involving many techniques of Natural Language Processing and learning models such as deep learning, support vector machine and random forest classifiers to distinguish MCI from control participants [48]. Speech is a major form of communication, and thus has great potential to monitor people with dementia. Such tools provide an automated and objective assessment of very early stage dementia that can be clinically monitored, thus enabling earlier diagnosis and timely interventions.

X. PSYCHOLOGICAL: SLEEP PATTERNS

Sleep cycles have long been linked to health. In the early 1990's research published by the American Academy of Neurology found particular anomalies, including disrupted cycles of sleep and waking and overall decrease in sleep have been found to be associated with dementia, though there was no correlation between the severity of the dementia and the amount of disrupt in the sleep cycle [49]. This was seconded by research finding fragmented sleep, frequent awakening a very little REM sleep being common in participants with Dementia. They also experienced tiredness and more frequently napped throughout the day [50]. Additionally, Caerphilly Cohort Study longitudinal study found that such disrupted sleep cycles was correlated with cognitive impairments.

XI. HEALTH RECORD MINING

A. Early Developments

Perhaps the broadest category discussed, health record mining looks at how data found in medical records can be used to inform vascular dementia diagnosis. This includes both text mining using Natural Language Processing techniques and numeric data record mining. One early case of record mining was a 15-year longitudinal look at blood pressure in relation to dementia onset. Finding showed that at age 70 there is a notable difference between control subjects and those who develop vascular dementia [51]. Interestingly, the Caerphilly Cohort Study also found blood pressure to be highly correlated

linked to quality of sleep and thus Dementia [51]. However, it's hard to make this information actionable, as it isn't specific enough to determine the cause high blood pressure to dementia.

B. Detecting Severity of Dementia

A study by Shankle et al. looks at estimating Clinical Dementia Rating Score, a dementia severity estimation scale, using machine learning and data found in Electronic Medical Records. The Clinical Dementia Rating requires a two-stage process to administer, only has 80% inter-rater reliability and costly and impractical. However, dementia severity is economically and clinically important, thus estimating the Clinical Dementia Rating has great benefit [52]. Using the C4.5, CART and Naïve Bayes machine learning algorithms, the model achieved 63 to 76 percent accuracy in predicting Clinical Dementia Rating with the Naïve Bayes model being the best. Of the four classes: normal, very mild, mild and moderately-severe, the mild class was difficult to classify and

co-mingled with very mild and moderately-severe. The other three classes achieved nearly 80 percent accuracy [52].

C. Differentiating Form of Dementia

A final area within the realm of data mining is to distinguish various forms of cognitive impairment. In a study by Manit et al. Alzheimer's and Vascular Dementia diagnoses were distinguished from Electronic Medical Record data [53]. This can be through text mining from written parts of the medical record or analysis of the numeric data, distinguishing features of patterns for particular types of dementia [54]. Using these techniques health record mining shows great potential in providing online, real-time differentiating diagnosis information for clinicians.

XII. SUMMARY OF RELATED WORKS

Below in Table 1, highlighted works from each of these above area are outlines; including their study design, data collection methods, tools used for data analytics and outcomes.

TABLE I. SUMMARY OF RELATED WORKS

| Publication, Author | Participant Description | Data Source (technological platform if applicable) | Data Analysis Techniques | Findings |
|--|---|---|---|--|
| "Distinctive cognitive profiles in Alzheimer's disease and subcortical vascular dementia." Graham, N. L., T. Emery, and J. R. Hodges [28] | Collected from 57 subjects -- 19 with subcortical vascular dementia, 19 with Alzheimer's disease, and 19 controls | Four cognitive domains: episodic memory, executive/attention functioning and visuospatial skills were assessed multiple using validated tests from each area. | Logistic regression analysis | The two groups could be discriminated with 89% accuracy on the basis of two tests, the WAIS logical memory – delayed recall test and a silhouette naming test using a logistic regression model. |
| "Visual Search in Patients with Subcortical Vascular Dementia: Short Fixations but Long Reaction Times" A. Rösler et al. [38] | A total of 18 subjects participated in the study -- 9 participants with Subcortical Vascular Dementia and 9 control subjects. | Reaction time, number of fixations and duration of fixations was measured from eye movement data recorded with a digital infrared eye tracker | Reaction time, number of fixations and duration of fixations were analyzed by 2-factorial ANOVA | In the Subcortical Vascular Dementia group, reaction time for the long distance array differed significantly from that for the arrays with shorter distances and from the reaction time for the control group at the same distance. |
| "Assessment of gait in subcortical vascular encephalopathy by computerized analysis: a cross-sectional and longitudinal study" H. Bänzner [42] | 119 patients with SVE (mean age 72±9.5 years; 61 men, 58 women). Of these, 39 were assessed in a longitudinal study with a mean interval of 26 months. | Using a on-body computerized gait analysis system, calculation of gait-lines representing the course of the force application point during foot heel- to-toe movement was used to score the participant's gait. | Statistical analysis evaluated the difference in scores using a t-test to measure the difference in score-defining variables. | The loss of a regular gait rhythm was noted in Subcortical Vascular Dementia. This is reflected by increased variability of all temporal variables calculating standard deviations of stance time, single and double support time. |
| "Abnormality of Gait as a Predictor of Non-Alzheimer's Dementia" J. Verghese [43] | The development of dementia was analyzed in a prospective study involving 422 subjects older than 75 years of age who lived in the community and did not have dementia at baseline. | Neurologic abnormalities affecting gait were diagnosed after clinical examination by board-certified neurologists. Gait was classified as unsteady if two or more abnormal features were present. | Cox proportional-hazards regression analysis to estimate hazard ratios with 95 percent confidence intervals for specific gait abnormalities, with adjustment for potentially confounding variables. | Unsteady gait was associated with an increased risk of vascular dementia (hazard ratio, 2.61; 95 percent confidence interval, 1.14 to 5.99), as was frontal gait (hazard ratio, 4.32; 95 percent confidence interval, 1.26 to 14.83) and hemiparetic gait (hazard ratio, 13.13; 95 percent confidence interval, 4.81 to 35.81) |

| Publication, Author | Participant Description | Data Source (technological platform if applicable) | Data Analysis Techniques | Findings |
|--|--|--|--|---|
| “Automatic speech analysis for the assessment of patients with predementia and Alzheimer’s disease” A. König. [47] | Healthy elderly control (HC) subjects and patients with MCI or AD were recorded while performing several short cognitive vocal tasks. | participant performed four spoken tasks during a regular consultation with a general practitioner while being recorded as a part of an ongoing research protocol. The tasks consisted of a counting backward task, a sentence repeating task, an image description task, and a verbal fluency task. | After recording, numerous vocal features were extracted from each spoken task. Analysis was language-independent -- speech recognition was not included, and only nonverbal features were targeted. | The accuracy of classification based on automatic audio analyses were as follows: between HCs and those with MCI, 79% 6 5%; between HCs and those with AD, 87% 6 3%; and between those with MCI and those with AD, 80% 6 5%, demonstrating its assessment utility. |
| “Predicting mild cognitive impairment from spontaneous spoken utterances” M. Asgari. [48] | 48 participants, including nine healthy controls, nine AD patients, and 30 frontotemporal lobar degeneration patients | Patients participated in social interaction sessions conducted using semi-structured conversations with trained interviewers for 30 minutes a day, 5 days a week for 6 weeks (i.e., 30 sessions). Words were then computationally modeled to extract meaning. Linguistic analysis of transcriptions began with grouping spoken words into 68 word subcategories. | Using computational analysis of narrative language samples, words are subcategorized into: (1) Linguistic Dimensions, (2) Psychological Processes, (3) Relativity, (4) Personal Concerns, and (5) Spoken Categories, and within these categories, further described by relational denotation, emotional connotation and denotation of time, space or motion. | To explore the effectiveness of different learning methods in distinguishing participants with MCI from those with intact cognition, we trained statistical models based on extracted linguistic features using two widely employed machine learning algorithms: (1) SVM and (2) random forest classifier (RFC) |
| “15-year longitudinal study of blood pressure and dementia” I. Skoog. [52] | As part of the Longitudinal Population Study of 70-year-olds in Göteborg, Sweden, we analysed blood pressure and dementia diagnostic. Participants continued followed up for 15 years. | The relation between blood pressure and the development of dementia was analyzed for each of the age intervals 70-75, 75-79, and 79-85 years in those non-demented at age 70. | Statistically differentiated. | Participants who developed dementia at age 79-85 had higher systolic blood pressure at age 70 and higher diastolic blood pressure at ages 70 and 75 than those who did not develop dementia. |
| “Simple Models for Estimating Dementia Severity Using Machine Learning” W. Shankle. [53] | The total sample consisted of the initial visits of 765 subjects ranging from normal to severely demented. | Patients received a complete diagnostic evaluation consisting of patient and caregiver interviews, general physical and neurological exam, two hours of cognitive testing. | Multiple Machine Learning algorithms were used to predict the Clinical Dementia Rating of a participant based on the provided data. | Participants were classified into one of four “levels” of Clinical Dementia Rating. Class accuracy ranged from 58-88% accuracy. Losses of accuracy were largely cases where the assigned classification was off by one class as ordered by severity. |

XIII. DISCUSSION

The discussed works use either statistical or machine learning techniques to differentiate cases of Mild Cognitive Impairment or Vascular Cognitive Impairment onset from control subjects. In the above studies, we see that half utilize some form of machine learning while the other half use ANOVA or regression based statistical analysis. In the case of machine learning techniques being used, it is common to see multiple methods being compared in effectiveness.

The choice of participant is another area of variation among the studies (see Fig. 2). Some studies were set up in a longitudinal fashion: in this realm majority of studies were set

up with the goal of detecting decline from a healthy to cognitively impaired state. However, one longitudinal study identified participants in early stages of Vascular Dementia onset in order to follow their progression. In the longitudinally designed studies comparisons are commonly drawn from the same individual but at different time periods. On the other end of the spectrum, some studies recruited participants previously diagnosed with some degree of Vascular Dementia. In this case, the study was largely focused on differentiating between the diagnosed and control groups, various levels or stages of dementia onset or between Vascular type Dementia and memory impairments caused from other forms of Dementia or Alzheimer’s Disease.

Finally, we look at the intentions of each of the discussed studies. The studies have been categorized into the following four classifications based on their outcomes: (1) distinguish between Vascular Dementia and other forms of Cognitive Impairment; (2) Distinguish between Vascular Dementia and Control group; (3) Describe the progression of particular Biometrics seen in Vascular Dementia patients; and (4) distinguish the severity of a particular case of Vascular Dementia. Utilizing these categories, the studied papers are distributed by achieved outcome type in Fig. 3.

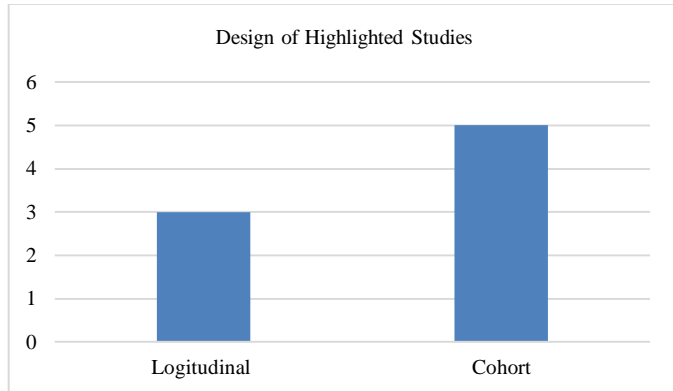


Fig. 2. Study Design of Highlighted Studies.

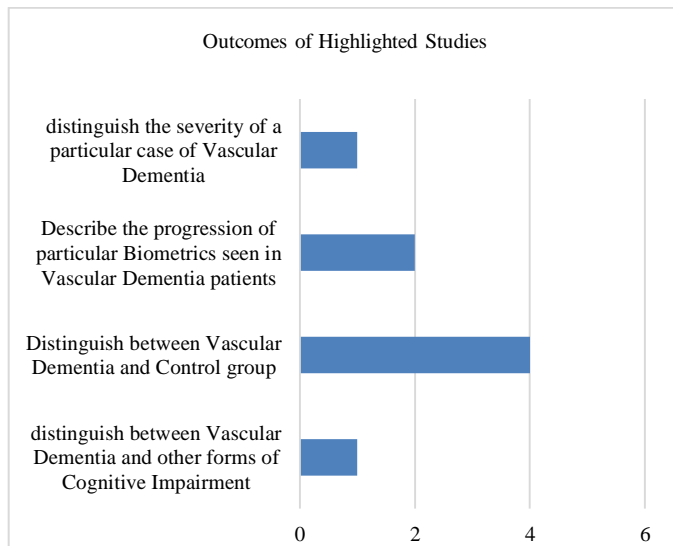


Fig. 3. Outcomes Achieved in Highlighted Studies.

XIV. CONCLUSION

An overarching limitation is that largely, research is validated findings against cognitive tests. This limits the findings to a timeframe when dementia is detectable using such assessments. In this sense, longitudinal studies could have important implications as they provide insight into the changes occurring before cognitive assessments can detect deterioration. We also noted that cognitive and behavioral assessments were some of the first assessments. In the realm of cognitive assessments, online and cognitive assessments have been incorporated into online assessments, visual tracking, and linguistic patterns. In terms of assessing behavioral changes, less has been integrated into the electronic environment. One

area that could be explored in Vascular Dementia onset is social network analysis. Through much of the discussed research, we see individual biometrics being correlated with the onset of vascular dementia. However, with the exception of Electronic Medical Record, most these methods looked a singular area of biometrics. Future work could investigate probabilistic models build using multiple data sources.

REFERENCES

- [1] Román, Gustavo. "Vascular dementia: a historical background." *International psychogeriatrics* 15.S1 (2003): 11-13.
- [2] Libon, David J., et al. "From Binswanger's disease to leukoaraiosis: What we have learned about subcortical vascular dementia." *The Clinical Neuropsychologist* 18.1 (2004): 83-100.
- [3] Pantoni, Leonardo, and Julio H. Garcia. "The significance of cerebral white matter abnormalities 100 years after Binswanger's report: a review." *Stroke* 26.7 (1995): 1293-1301.
- [4] Bonnici-Mallia, Anne M., Christopher Barbara, and Rahul Rao. "Vascular cognitive impairment and vascular dementia." *InnovAiT* (2018):1755738018760649.
- [5] V. J. A. Verlinden, J. N. van der Geest, R. F. A. G. de Bruijn, A. Hofman, P. J. Koudstaal, and M. A. Ikram. "Trajectories of decline in cognition and daily functioning in preclinical dementia". *Alzheimer's & Dementia* 12.2 (Feb. 1, 2016), pp. 144-153. issn: 1552-5260. Doi: 10.1016/j.jalz.2015.
- [6] Allali, Gilles, Marian Van Der Meulen, and Frédéric Assal. "Gait and cognition: The impact of executive function." *Swiss Archives of Neurology and Psychiatry* 161.6 (2010): 195-199.
- [7] S. Gauthier, B. Reisberg, M. Zaudig, R. C. Petersen, K. Ritchie, K. Broich, S. Belleville, H. Brodaty, D. Bennett, H. Chertkow, J. L. Cummings, M.deLeon, H. Feldman, M. Ganguli, H. Hampel, P. Scheltens, M. C. Tierney, P.Whitehouse, and B. Winblad. "Mild cognitive impairment". *The Lancet* 367.9518 (Apr. 15, 2006), pp. 1262-1270. issn: 0140-6736. Doi: 10.1016/S0140-6736(06)68542-5.
- [8] B. C. Stephan, F. E. Matthews, K.-T. Khaw, C. Dufouil, and C. Brayne. "Beyond mild cognitive impairment: vascular cognitive impairment, no dementia(VCIND)". *Alzheimer's Research & Therapy* 1.1 (July 9, 2009), p. 4. issn: 1758-9193. doi: 10.1186/alzrt4.
- [9] D. S. Knopman. "The initial recognition and diagnosis of dementia". *The American journal of medicine* 104.4 (1998), 2S-12.
- [10] García, PL Rodríguez, and D. Rodríguez García. "Diagnosis of vascular cognitive impairment and its main categories." *Neurología (English Edition)* 30.4 (2015): 223-239.
- [11] Chandola, Varun, Arindam Banerjee, and Vipin Kumar. "Anomaly detection: A survey." *ACM computing surveys (CSUR)* 41.3 (2009): 15.
- [12] DasGupta, Dipankar. "An overview of artificial immune systems and their applications." *Artificial immune systems and their applications*. Springer, Berlin, Heidelberg, 1993. 3-21.
- [13] Jain, Anil K., Arun Ross, and Salil Prabhakar. "An introduction to biometric recognition." *IEEE Transactions on circuits and systems for video technology* 14.1 (2004): 4-20.
- [14] Marshall, John C., and Freda Newcombe. "Patterns of paralexia: A psycholinguistic approach." *Journal of psycholinguistic research* 2.3 (1973): 175-199.
- [15] Reed, Stephen K. "Pattern recognition and categorization." *Cognitive psychology* 3.3 (1972): 382-407.
- [16] Valiant, Leslie G. "A theory of the learnable." *Communications of the ACM* 27.11 (1984): 1134-1142.
- [17] Lazer, David, et al. "Life in the network: the coming age of computational social science." *Science (New York, NY)* 323.5915 (2009): 721.
- [18] Hall, David L., and James Llinas. "An introduction to multisensor data fusion." *Proceedings of the IEEE* 85.1 (1997): 6-23.
- [19] MacKay, David JC, and David JC Mac Kay. *Information theory, inference and learning algorithms*. Cambridge university press, 2003.

- [20] Liu, Huan, and Hiroshi Motoda, eds. Feature extraction, construction and selection: A data mining perspective. Vol. 453. Springer Science & Business Media, 1998.
- [21] Stillou, S., et al. "Mining association rules from clinical databases: an intelligent diagnostic process in healthcare." *Studies in health technology and informatics* 2 (2001): 1399-1403.
- [22] Hillestad, Richard, et al. "Can electronic medical record systems transform health care? Potential health benefits, savings, and costs." *Health affairs* 24.5 (2005): 1103-1117.
- [23] Jain, Anil K., Robert PW Duin, and Jianchang Mao. "Statistical pattern recognition: A review." *IEEE Transactions on pattern analysis and machine intelligence* 22.1 (2000): 4-37.
- [24] Kononenko, Igor. "Machine learning for medical diagnosis: history, state of the art and perspective." *Artificial Intelligence in medicine* 23.1 (2001): 89-109.
- [25] Yoo, Illhoi, et al. "Data mining in healthcare and biomedicine: a survey of the literature." *Journal of medical systems* 36.4 (2012): 2431-2448.
- [26] Román, Gustavo C. "A historical review of the concept of vascular dementia: lessons from the past for the future." *Alzheimer disease and associated disorders* 13.3 (1999): S4.
- [27] Nyenhuis, David L., and Philip B. Gorelick. "Vascular dementia: a contemporary review of epidemiology, diagnosis, prevention, and treatment." *Journal of the American Geriatrics Society* 46.11 (1998): 1437-1448.
- [28] Graham, N. L., T. Emery, and J. R. Hodges. "Distinctive cognitive profiles in Alzheimer's disease and subcortical vascular dementia." *Journal of Neurology, Neurosurgery & Psychiatry* 75.1 (2004): 61-71.
- [29] Frank, Rowena M., and Gerard J. Byrne. "The clinical utility of the Hopkins Verbal Learning Test as a screening test for mild dementia." *International Journal of Geriatric Psychiatry* 15.4 (2000): 317-324.
- [30] Dong, YanHong, et al. "The Montreal Cognitive Assessment is superior to the Mini-Mental State Examination in detecting patients at higher risk of dementia." *International Psychogeriatrics* 24.11 (2012): 1749-1755.
- [31] Loring, David W. "The Wechsler memory scale-revised, or the Wechsler memory scale-revisited?." *The Clinical Neuropsychologist* 3.1 (1989): 59-69.
- [32] Grober, Ellen, et al. "Free and cued selective reminding distinguishes Alzheimer's disease from vascular dementia." *Journal of the American Geriatrics Society* 56.5 (2008): 944-946.
- [33] Davis, Kelly L., et al. "Error analysis of the nine-word California Verbal Learning Test (CVLT-9) among older adults with and without dementia." *The Clinical Neuropsychologist* 16.1 (2002): 81-89.
- [34] Gualtieri, C. Thomas. "Dementia screening using computerized tests." *Journal of Insurance Medicine-New York then Denver--* 36 (2004): 213-227.
- [35] Román, Gustavo. "Vascular dementia: a historical background." *International psychogeriatrics* 15.S1 (2003): 11-13.
- [36] Lyketsos, Constantine G., et al. "Mental and behavioral disturbances in dementia: findings from the Cache County Study on Memory in Aging." *American Journal of Psychiatry* 157.5 (2000): 708-714.
- [37] Nyenhuis, David L., et al. "The pattern of neuropsychological deficits in vascular cognitive impairment-no dementia (vascular CIND)." *The Clinical Neuropsychologist* 18.1 (2004): 41-49.
- [38] Rösler, A., et al. "Visual search in patients with subcortical vascular dementia: Short fixations but long reaction times." *Dementia and geriatric cognitive disorders* 20.6 (2005): 375-380.
- [39] Kokkinakis, Dimitrios, et al. "Data collection from persons with mild forms of cognitive impairment and healthy controls—infrastructure for classification and prediction of dementia." *Proceedings of the 21st Nordic Conference on Computational Linguistics*. 2017.
- [40] Crutcher, Michael D., et al. "Eye tracking during a visual paired comparison task as a predictor of early dementia." *American Journal of Alzheimer's Disease & Other Dementias* 24.3 (2009): 258-266.
- [41] Rösler, A., et al. "Effects of arousing emotional scenes on the distribution of visuospatial attention: Changes with aging and early subcortical vascular dementia." *Journal of the neurological sciences* 229 (2005): 109-116.
- [42] Bänzner, H., et al. "Assessment of gait in subcortical vascular encephalopathy by computerized analysis: a cross-sectional and longitudinal study." *Journal of neurology* 247.11 (2000): 841-849.
- [43] Verghese, Joe, et al. "Abnormality of gait as a predictor of non-Alzheimer's dementia." *New England Journal of Medicine* 347.22 (2002): 1761-1768.
- [44] Hausdorff, Jeffrey M., and Aron S. Buchman. "What links gait speed and MCI with dementia? A fresh look at the association between motor and cognitive function." (2013): 409-411.
- [45] Román, Gustavo C., and Donald R. Royall. "Executive control function: a rational basis for the diagnosis of vascular dementia." *Alzheimer disease and associated disorders* (1999).
- [46] Fitten, L. Jaime, et al. "Alzheimer and vascular dementias and driving: a prospective road and laboratory study." *Jama* 273.17 (1995): 1360-1365.
- [47] König, Alexandra, et al. "Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease." *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* 1.1 (2015): 112-124.
- [48] Asgari, Meysam, Jeffrey Kaye, and Hiroko Dodge. "Predicting mild cognitive impairment from spontaneous spoken utterances." *Alzheimer's & Dementia: Translational Research & Clinical Interventions* 3.2 (2017): 219-228.
- [49] Aharon-Peretz, J., et al. "Sleep-wake cycles in multi-infarct dementia and dementia of the Alzheimer type." *Neurology* 41.10 (1991): 1616-1616.
- [50] Elwood, Peter C., et al. "Sleep disturbance and daytime sleepiness predict vascular dementia." *Journal of Epidemiology & Community Health* 65.9 (2011): 820-824.
- [51] Prinz, Patricia N., et al. "Changes in the sleep and waking EEGs of nondemented and demented elderly subjects." *Journal of the American Geriatrics Society* 30.2 (1982): 86-92.
- [52] Skoog, Ingmar, et al. "15-year longitudinal study of blood pressure and dementia." *The Lancet* 347.9009 (1996): 1141-1145.
- [53] Shankle, William Rodman, et al. "Simple models for estimating dementia severity using machine learning." *MedInfo* 98 (1998).
- [54] Mani, Subramani, et al. "Differential diagnosis of dementia: A knowledge discovery and data mining (KDD) approach." *Proceedings of the AMIA Annual Fall Symposium*. American Medical Informatics Association, 1997.