Phenotyping Cognitive Impairment using Graphomotor and Latency Features in Digital Clock Drawing Test *

Anis Davoudi, Catherine Dion, Shawna Amini, David J. Libon, Patrick J. Tighe, Catherine C. Price, Parisa Rashidi

***Abstract*— The Clock Drawing Test, where the participant is asked to draw a clock from memory and copy a model clock, is widely used for screening of cognitive impairment. The digital version of the clock test, the digital clock drawing test (dCDT), employs accelerometer and pressure sensors of a digital pen to capture time and pressure information from a participant's performance in a granular digital format. While visual features of the clock drawing test have previously been studied, little is known about the relationship between demographic and cognitive impairment characteristics with dCDT latency and graphomotor features. Here, we examine dCDT feature clusters with respect to sociodemographic and cognitive impairment outcomes. Our results show that the clusters are not significantly different in terms of age and gender, but did significantly differ in terms of education, Mini-Mental State Exam scores, and cognitive impairment diagnoses.**

***Clinical Relevance*— This study shows that features extracted from digital clock drawings can provide important information regarding cognitive reserve and cognitive impairments.**

I. Introduction

The 2019 American College of Surgeons Geriatric Surgery Verification Program recommends perioperative cognitive screening for older adults [1]. Early identification of cognitive impairment in perioperative medicine is essential to limit the risks of developing delirium and post-operative cognitive deficits, particularly in older populations [1]. The clock drawing test (CDT) is easy to use and a validated perioperative cognitive screening tool (Figure I) [2-5]. Participants are asked to draw the clock face to show a specific time, usually 11:10 (c*ommand* condition) followed by copying a model of the clock (c*opy* condition). Easily administered at bedside, the CDT is quick and efficient, making it well-suited for perioperative settings. Nonetheless, there are several drawbacks to using the conventional CDT. First, scoring relies on subjective judgment of the final product only. Second temporal and motor information during performance cannot be recorded using the traditional pen and paper task. The digital clock drawing test (dCDT), which uses digital pen technology, makes it possible to capture timing, pressure of the pen while drawing, and other features (Figure II) [6, 7]. Pen positioning on paper is captured with ±0.005 spatial and 12 milliseconds temporal accuracy. This newly available information also yields data assessing latency and graphomotor variables of interest providing additional insight into the neurocognitive domains underlying test performance [8]. Several recent studies have shown the feasibility of using dCDT test to study putative cognitive impairments in patient groups [6, 9]. However, the relationship between demographic and cognitive impairment characteristics with dCDT latency and graphomotor features remain unknown. Here, we compare the distribution of demographic and cognitive impairment characteristics by group by clustering graphomotor and latency data captured from dCDT.

II. Methods

A. Dataset

We collected our data from two cohorts of participants diagnosed with dementia or Parkinson disease (PD), and a control cohort of healthy participants. Participants were asked to perform the clock drawing task using aforementioned digital pen technology to capture the pen movement. Only command condition dCDT data was considered for this project. We also obtained their demographics and Mini-Mental State Examination (MMSE) [10] scores.

B. Analysis

We extracted 29 features that we identified from relevant literature (Table I) [6, 7, 9, 11, 12]. To impute continuous features, we use Multiple Imputation using Chained Equations (MICE) algorithm [13], using example values from participants with similar profiles with respect to other variables. Next, we used the Uniform Manifold Approximation and Projection (UMAP) [14] method for dimensionality reduction and transformed data into 2 dimensions.

We used the K-means clustering algorithm to detect clusters using the extracted dCDT features. We also used Hopkins statistics to check for clustering tendency of the data and used the average Silhouette metric for detection of the optimal number of clusters [15, 16]. The silhouette metric evaluates how well the data points are matched to other points in their cluster and how different are from points in other clusters. The

*This work is supported by NSF CAREER 1750192 (PR), and NIH/NIBIB 1R21EB027344 (PR), R01 NR014181, R01NS082386, NSF 13-543, R01 AG055337-02.

A. Davoudi is with the Biomedical Engineering Department, University of Florida, Gainesville, FL 32611 USA (email: anisdavoudi@ufl.edu)

C. Dion is with the Department of Department of Clinical and Health Psychology, University of Florida, Gainesville, FL 32611 USA (email: cdion2@phhp.ufl.edu)

S. Amini is with the Department of Department of Clinical and Health Psychology, University of Florida, Gainesville, FL 32611 USA (email: samini@anest.ufl.edu)

D. J. Libon is with the Department of Geriatric, Rowan University, Glassboro, NJ 08028 USA (email: libon@rowan.edu)

P. J. Tighe is with the Department of Anesthesiology, University of Florida, Gainesville, FL 32611 USA (email: ptighe@anest.ufl.edu)

C. Price is with the Department of Clinical and Health Psychology, University of Florida, Gainesville, FL 32611 USA (email: cep23@phhp.ufl.edu)

P. Rashidi is with the Department of Biomedical Engineering, University of Florida, Gainesville, FL 32611 USA (email: parisa.rashidi@ufl.edu)

Figure I Example of Copy and Command clockfaces in the dCDT test.

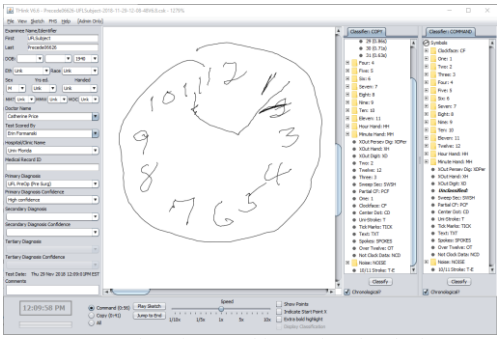

Figure II Program interface used in scoring the clocks.

average Silhouette method calculates the Silhouette metric for different numbers of clusters and determines the optimal number of clusters as the one with the highest average Silhouette. Next, we compared the profiles of the participants in the three detected clusters in terms of demographics and diagnosis. Kruskal-Wallis and Chi-square tests were sued to compare the distribution of studied variables. The analyses were performed using R 3.6.2.

III. Results

Our dataset consists of 316 participants (Table I): 166 (52.5%) healthy Control participants, 70 (22.2%) Dementia patients, and 80 (25.3%) non-demented participants with Parkinson's disease (PD). Age, gender, education and MMSE scores were significantly different between the three classes. Dementia patients were generally older (73.9±13.1 years compared with 64.9±13.3 control participants and 68.8±6.4 for participants with PD). Dementia patients also had fewer years of education (12.8±2.7 compared with 15.9±2.6 for control participants and 16.6±2.5 for PD participants, and lower MMSE scores (23.3±3.2 compared with 28.6±1.3 for control participants and 28.5±1.3 for PD participants). Distribution of gender was similar between dementia patients and control participants, while there were fewer female Parkinson patients.

Figure III shows the resulting two dimensions after using UMAP for dimensionality reduction. Silhouette metric and elbow method were used for determining the optimal number of clusters using K-means clustering algorithm. Both Silhouette metric found three clusters as the optimal number of clusters detected in the dataset (Figure IV). Visual inspection of the resulting clusters in two dimensions also shows separation between the clusters (Figure V). Next, we compared the demographic variables of age, gender, and education between the three clusters. Education, unlike age and gender, was significantly different between the three clusters (p-values: 1.8 e-6, 0.377, and 0.735, respectively; Figure VI and Table III). To compare the clusters in terms of cognitive performance, we compared the distribution of MMSE scores and cognitive impairment diagnoses among the three clusters, both being different (p-values: 1.4 e-5 and 3.9 e-6, respectively, Table III, Figure VII, VIII).

IV. Discussion

In this study, we examined the distribution of demographic variables (age, gender, and education) in clusters formed using graphomotor and related features extracted from digital clock drawing in participants with different clinical diagnoses. Results show that while age and gender are not significantly different between the clusters, education, MMSE scores, and cognitive diagnoses were significantly different between the participants in these clusters. Results suggest education can be considered as an index of cognitive reserve and was significantly different between the three cohorts, with dementia patients having significantly fewer years of education compared to participants with PD and healthy controls. Our work had several limitations. The cohort of

Table I Features extracted from pen movement data.

total completion time	Average pressure	Minute-hand length/hour-hand length
time in air	Standard deviation of pressure	Pre-first-hand latency
Time on paper	Average velocity	Post-clockface latency
strokes per minute	Standard deviation of velocity	Anchoring
total number of strokes	Average velocity/average pressure	Inter-digit latency
pen-up length	Clockface area	Inter-digit interval count
Pen-down length	Minute-hand length	At least one digit missing
Pen-up length/pen-down length	Hour-hand length	Any digit over 12
Any repeated digit	Average digit height	Average digit width

TABLE II DISTRIBUTION OF VARIABLES AMONG DIAGNOSIS GROUPS.

Variable	All (N=316)	Control (N=166) Mean (SD)	Dementia (N=70) Mean (SD)	Parkinson's disease (N=80) Mean (SD)	p-value
Age	67.9 (12.4)	64.9 (13.3)	73.9 (13.1)	68.8 (6.4)	**2.4 e-7**
Gender, Female N (%)	159 (50.3)	91 (54.8)	39 (55.7)	29 (36.3)	**0.014**
Education	15.4 (3.0)	15.9 (2.6)	12.8 (2.7)	16.6 (2.5)	**3.7 e-16**
MMSE	27.4 (2.9)	28.6 (1.3)	23.3 (3.2)	28.5 (1.3)	**<2.2 e-16**

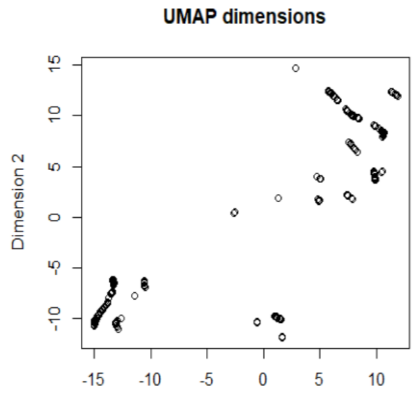

Figure III UMAP representation of the data in two dimensions.

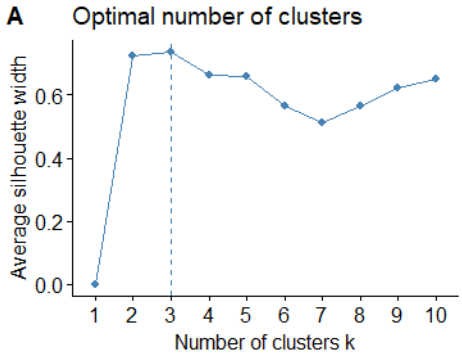

Figure IV Optimal number of clusters using K-means clustering algorithm and average Silhouette.

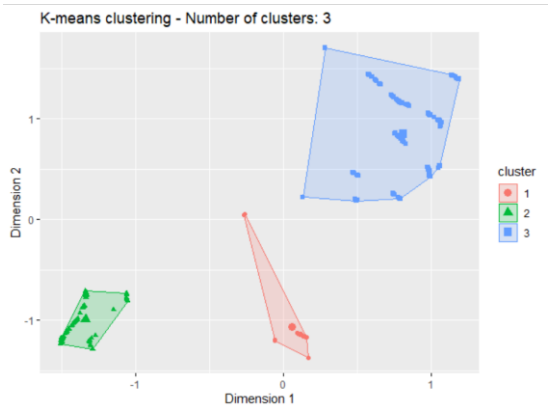

Figure V Three clusters using K-means clustering algorithm

patients diagnosed with Parkinson's disease were all in early stages of the disease. As a result, the demographics and MMSE scores of the healthy control participants and Parkinson's disease patients were not significantly different. Future work needs to examine the differentiation power of clustering algorithms with more general PD populations with varying levels of disease duration. Moreover, different cognitive phenotypes of Parkinson's disease (i.e. cognitively well, low memory, and low executive function) as well as different subtypes of dementia need to be further examined for stratified clustering and analysis, since different cognitive domains might be affected differently in each subtype. Another limitation was that the variety in data collection settings might affect the pressure data captured by the pen, and needs to be

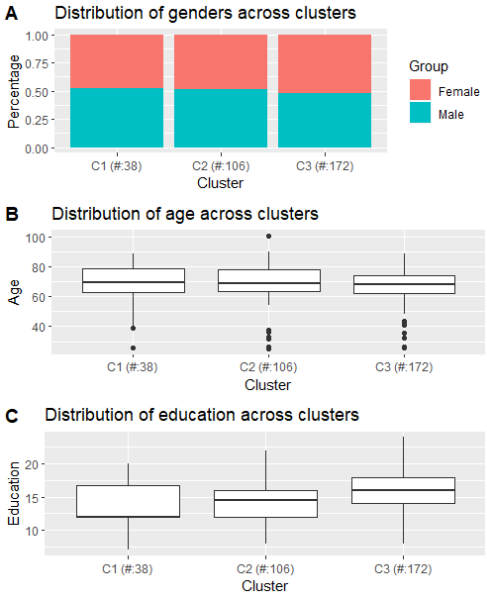

Figure VI Distribution of demographics -gender (A), age (B), and education (C) - among the detected clusters.

further investigated. In our future work, we plan to apply our methodology to a larger cohort with more detailed cognitive diagnoses and neuropsychology performance variables to better understand how the dCDT may be used for cognitive

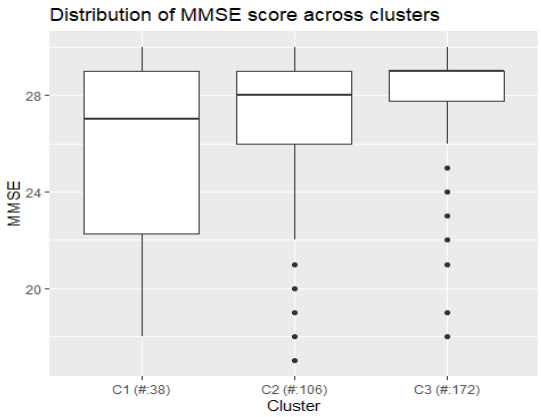

Figure VII Distribution of MMSE scores across the clusters.

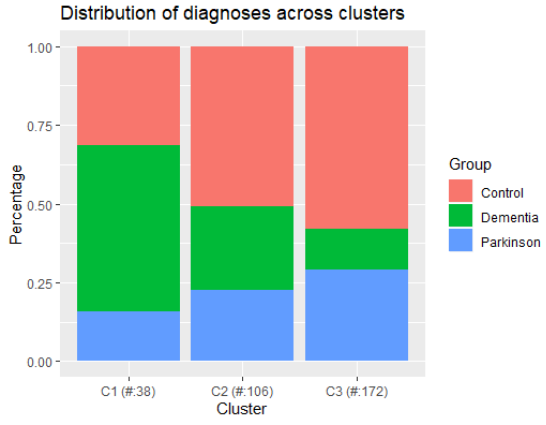

Figure VIII Distribution of diagnosis among the clusters.

impairment screening. Additionally, we plan to investigate the enhancement of classification models by using multimodal models using clock images in addition to the pen movement time series data.

V. Acknowledgment

We acknowledge Dr. Dana Penney at Lahey Hospital and Medical Center, Burlington, Massachusetts, and Dr. Randal Davis at Massachusetts Institute of Technology for their expertise and help in using the digital Clock Drawing Test Platform.

TABLE III DISTRIBUTION OF VARIABLES AMONG DETECTED CLUSTERS.

Variable	All (N=316)	Cluster 1 (N=38) Mean (SD)	Cluster 2 (N=106) Mean (SD)	Cluster 3 (N=172) Mean (SD)	p-value
Age	67.9 (12.4)	67.9 (14.5)	68.5 (14.1)	67.5 (10.7)	0.377
Gender, Female N (%)	159 (50.3)	18 (47.4)	51 (48.1)	90 (52.3)	0.735
Education	15.4 (3.0)	14.0 (3.1)	14.6 (2.9)	16.1 (2.8)	**1.8 e-6**
MMSE	27.4 (2.9)	25.6 (3.7)	27.0 (3.0)	28.1 (2.3)	**1.4 e-5**
Diagnosis	Control: 166 (52.5) Dementia: 70 (22.2) Parkinson's disease: 80 (25.3)	Control: 12 (31.6) Dementia: 20 (52.6) Parkinson's disease: 6 (15.8)	Control: 54 (50.9) Dementia: 28 (26.4) Parkinson's disease: 24 (22.6)	Control: 100 (58.1) Dementia: 22 (12.8) Parkinson's disease: 50 (29.1)	**3.9 e-6**

Bibliography

[1] "Optimal Resources for Geriatric Surgery: 2019 Standards." American College of Surgeons. https://www.facs.org/-/media/files/quality-programs/geriatric/geriatricsv_standards.ashx (accessed 1/13/2020, 2020).

[2] S. Amini *et al.*, "Feasibility and rationale for incorporating frailty and cognitive screening protocols in a preoperative anesthesia clinic," *Anesthesia and analgesia,* vol. 129, no. 3, p. 830, 2019.

[3] D. J. Culley *et al.*, "Preoperative cognitive stratification of older elective surgical patients: a cross-sectional study," *Anesthesia and analgesia,* vol. 123, no. 1, p. 186, 2016.

[4] M. Freedman, *Clock drawing: A neuropsychological analysis*. Oxford University Press, USA, 1994.

[5] L. J. Grande *et al.*, "Clock drawing: Standing the test of time," *The Boston process approach to neuropsychological assessment. A practitioner's guide,* pp. 229-248, 2013.

[6] W. Souillard-Mandar *et al.*, "Learning Classification Models of Cognitive Conditions from Subtle Behaviors in the Digital Clock Drawing Test," (in eng), *Mach Learn,* vol. 102, no. 3, pp. 393-441, Mar 2016, doi: 10.1007/s10994-015-5529-5.

[7] R. Davis, D. J. Libon, R. Au, D. Pitman, and D. L. Penney, "THink: Inferring Cognitive Status from Subtle Behaviors," (in eng), *Proc Conf AAAI Artif Intell,* vol. 2014, pp. 2898-2905, Jul 2014.

[8] C. Dion *et al.*, "Cognitive Correlates of Digital Clock Drawing Metrics in Older Adults with and without Mild Cognitive Impairment," (in eng), *J Alzheimers Dis,* Apr 1 2020, doi: 10.3233/jad-191089.

[9] S. Muller *et al.*, "Diagnostic value of digital clock drawing test in comparison with CERAD neuropsychological battery total score for discrimination of patients in the early course of Alzheimer's disease from healthy individuals," (in eng), *Sci Rep,* vol. 9, no. 1, p. 3543, Mar 5 2019, doi: 10.1038/s41598-019-40010-0.

[10] L. Kurlowicz and M. Wallace, "The mini-mental state examination (MMSE)," *Journal of gerontological nursing,* vol. 25, no. 5, pp. 8-9, 1999.

[11] L. P. Hizel *et al.*, "Clock Drawing Performance Slows for Older Adults After Total Knee Replacement Surgery," (in eng), *Anesth Analg,* vol. 129, no. 1, pp. 212-219, Jul 2019, doi: 10.1213/ane.0000000000003735.

[12] M. Lamar *et al.*, "Cognitive and connectome properties detectable through individual differences in graphomotor organization," (in eng), *Neuropsychologia,* vol. 85, pp. 301-9, May 2016, doi: 10.1016/j.neuropsychologia.2016.03.034.

[13] M. J. Azur, E. A. Stuart, C. Frangakis, and P. J. Leaf, "Multiple imputation by chained equations: what is it and how does it work?," *International journal of methods in psychiatric research,* vol. 20, no. 1, pp. 40-49, 2011.

[14] L. McInnes, J. Healy, and J. Melville, "Umap: Uniform manifold approximation and projection for dimension reduction," *arXiv preprint arXiv:1802.03426,* 2018.

[15] L. Kaufman and P. J. Rousseeuw, *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons, 2009.

[16] B. Hopkins and J. G. Skellam, "A new method for determining the type of distribution of plant individuals," *Annals of Botany,* vol. 18, no. 2, pp. 213-227, 1954.