Supplementary Material

Combining Multimodal Behavioral Data of Gait, Speech, and Drawing for Classification of Alzheimer's Disease and Mild Cognitive Impairment

Supplemental Methods

Structural magnetic resonance imaging

All participants were administered structural magnetic resonance imaging (MRI) scans with 1.5T, T1-weighted images and a 3D gradient-echo sequence with the following parameters: sagittal orientation with 1.2-mm-thick sections; time repetition/time echo 2400/3.52 milliseconds; flip angle 8°; field of view 240×240. We expressed the severity of medial temporal lobe atrophy as a Z score relative to cognitively healthy adults by using a standalone, voxel-based specific regional analysis system for Alzheimer disease (AD) [1].

Gait data collection and feature extraction

Gait data were concurrently recorded using an eight-camera OptiTrack Flex 13 motion capture system, sampled at 120 Hz using OptiTrack Motive software 2.1.0 Beta 1 (NaturalPoint, Inc, Corvallis, OR, USA). Fifty reflective, spherical markers were applied to defined anatomical landmarks in accordance with the marker setup of the OptiTrack Motive software. To discard the increase and decrease speed effect, the first and last two meters were excluded from analyses.

We extracted a total of 35 gait features associated with pace, rhythm, variability, left-right asymmetry, and postural control based on previous studies on gait characteristics related to cognitive impairment and AD [2–6] as well as functioning changes in older adults related to fall risks [7–10] and frailty [11,12]. Specifically, 14 features were related to steps and strides: mean and variability of step length, step time, step angle, step width, stride length, and stride time as well as absolute differences between left and right step time and length. We calculated gait speed, acceleration, and jerk by using the trajectory of the pelvis marker on the waist back. We extracted 15 features including the mean, variability, maximum, and root mean square of gait speed, acceleration, and jerk. In addition, the first and second peaks computed on the basis of the unbiased autocorrelation coefficients (i.e., Ad1 and Ad2) [13] were used as features to measure left-right asymmetry between steps and the variability of consecutive strides. Mediolateral fluctuation of the pelvis marker was also used as a gait feature related to postural control. We calculated the speed of the heel markers relative to the pelvis marker for measuring foot swing speed and used three features related to the mean, variability, and left-right differences of foot swing speed. Finally, maximum toe clearance, toe off angle, and heal strike angle were used as features related to postural control.

Speech data collection and feature extraction

For the speech data collection, participants sat down in front of the tablet and answered questions presented by a voice-based application on the tablet. The tablet showed a screen indicating whether it was speaking or listening. We used an iPad Air 2 and recorded voice responses by using the iPad's internal microphone (core audio format, 44,100 Hz, 16-bit). The audio data were manually transcribed. Filler words such as “um” and “uh” were annotated manually and were considered as pauses used for extracting pause-related features.

From the speech data of each participant, we extracted a total of 84 speech features used in previous studies on detecting patients with MCI or AD [14–21]. Speech features consisted of 58 acoustic features, 15 prosodic features, and 11 linguistic features. We extracted features from each task unless otherwise stated.

The acoustic and prosodic features were extracted from audio data. The acoustic features consisted of three feature types related to shimmer, jitter, and Mel-frequency cepstral coefficients (MFCCs). Jitter and shimmer features are cycle-to-cycle variations of fundamental frequency and amplitude, commonly used to measure pathological voice quality [22]. For jitter, we used local jitter calculated as the average absolute difference between consecutive periods divided by the average period [22]. For shimmer, we used local shimmer as the average absolute difference between the amplitudes of consecutive periods divided by the average amplitude [22]. MFCCs are spectral features characterizing the frequency distribution of a speech signal in specific time instance information and designed to take into account the response properties of the human auditory system [23]. We used the mean, variance, skewness, and kurtosis of the first 12 MFCCs during spontaneous speech, that is, the picture description task. The prosodic features included phoneme rate, pitch variability, and proportion of pause duration. For estimating pitch, we used fundamental frequency. We used the following audio-processing libraries in Python (version 3.8): librosa (version 0.8.0 [24]) for calculating MFCCs and Signal_Analysis (version 0.1.26 [25]) for calculating fundamental frequency.

The linguistic features were extracted from manually transcribed text data. The linguistic features were the proportion of mistakes in the counting backwards and subtraction tasks, number of correct answers in the phonemic and semantic verbal fluency tasks, and Honoré's statistics (HS) [26] for measuring vocabulary richness and six features related to the number of information units in the picture description task. HS gives particular importance to unique words used only once and is calculated by the following equation: $\mathrm{HS} = 100 \log N/(1 - V_1/V)$, where N is the total number of words, V is the number of unique words, and V_1 is the number of words spoken only once. The number of information units was obtained by counting the number of unique pre-defined entities relevant to the picture. Whether participants mentioned each entity was determined by manually annotating entities to words in the transcribed text data. For example, words such as boy, son, or brother were all annotated as the entity “boy.” We used 23 entities, or information units, in four categories based on previous studies [27,28]: subjects (mother, boy, and girl), places (kitchen and exterior seen through the window), objects (faucet, water, sink and counter, floor, plate and dishes on the counter, dishcloth, cookies, jar, cabinet, stool, window, and curtain), and actions (boy taking the cookie, boy or stool falling, mother drying or washing the dishes, water overflowing, girl asking for a cookie, and mother unconcerned by the water overflowing or children stealing cookies). We used the number of information units for each category and the total number of all information units as well as the total number of information units normalized by the length of the speech. For tokenizing and lemmatizing the transcribed text data, we used the Japanese morphological analyzer Janome (version 0.4.1 [29]) in Python (version 3.8).

Drawing data collection and feature extraction

We collected drawing data using a Wacom Cintiq Pro 16 (sampling rate: 180 Hz, pen pressure levels: 8192). We then extracted a total of 60 drawing features based on previous studies on drawing characteristics related to cognitive impairment [30,31], MCI or AD [30,32–36], and other neurodegenerative diseases related to physical functioning changes such as Parkinson disease [37,38]. They consisted of 23 kinematic features, 5 pressure-related features, 20 time-related features, and 12 trail-making-test-specific features. We extracted features from each task unless otherwise stated. The kinematic features included the mean, variability, and maximum of the drawing speed, mean drawing acceleration, total stroke length in the sentence task, and total stroke length normalized by number of correct edges in the trail-making tasks. For the pressure-related features, we used writing pressure variability. The time-related features included the average and maximum of pause duration between drawings (i.e., between strokes or within a stroke), pause duration between drawings normalized by total stroke length, and task duration normalized by total stroke length. The trail-making-test-specific features included the number of correct and incorrect edges, number of lifts, mean and maximum of time between nodes, and task duration normalized by the number of answered edges.

Preprocessing and parameters for classification models

The number of input features from individual behavioral modalities was set to 35, the smallest number of gait features among the modalities, to be the same as the number of features from individual modalities. Input features of drawing and speech features were selected on the basis of an area under receiver operating characteristic curve (AuROC). As for missing values, multivariate imputation by chained equations [39] was carried out using all non-missing input features of the same behavioral modality. The parameters that we studied were as follows: the number of neighbors (search range: 1, 2, 3, 5, 10, 15, 20) for the k-nearest neighbors; the number of trees (10, 50, 100), maximum depth of trees (2, 3, 4, 5), class weights (None, balanced, balanced_subsample) for random forest; kernel functions (linear and radial basis function), penalty parameter (1, 10, 20, 50, 70, 90, 100, 110, 150, 200), the parameter associated with the width of the radial basis function kernel (1, 0.5, 0.1, 0.05, 0.01, 0.005, 0.001, 0.0005, 0.0001, $1/N_{\mathrm{features}}$), and class weights (None, balanced) for the support vector machine. N_{features} is the number of input features. The parameters were tuned using ten-fold cross validation. We used algorithms implemented using the Python package scikit-learn (version 0.23.2) and all other parameters were kept at the default values.

Supplemental Results

Missing behavioral data

Four AD patients did not complete the five speech tasks. One AD patient did not perform the picture description task, one did not perform the counting backwards task, and one performed only the picture description task because they could not follow the instructions. The remaining patient did not perform the counting backwards and subtraction tasks due to the participant's verbal refusal.

REFERENCES

[1] Matsuda H, Mizumura S, Nemoto K, Yamashita F, Imabayashi E, Sato N, Asada T (2012) Automatic voxel-based morphometry of structural MRI by SPM8 plus diffeomorphic anatomic registration through exponentiated lie algebra improves the diagnosis of probable Alzheimer disease. *AJNR Am J Neuroradiol* **33**, 1109–1114.

[2] Beauchet O, Allali G, Montero-Odasso M, Sejdić E, Fantino B, Annweiler C (2014) Motor phenotype of decline in cognitive performance among community-dwellers without dementia: population-based study and meta-analysis. *PLoS One.* **9**, e99318.

[3] Mielke MM, Roberts RO, Savica R, Cha R, Drubach DI, Christianson T, Pankratz VS, Geda YE, Machulda MM, Ivnik RJ, Knopman DS, Boeve BF, Rocca WA, Petersen RC (2013) Assessing the temporal relationship between cognition and gait: slow gait predicts cognitive decline in the Mayo Clinic Study of Aging. *J Gerontol A Biol Sci Med Sci* **68**, 929–937.

[4] Mc Ardle R, Galna B, Donaghy P, Thomas A, Rochester L. Do Alzheimer's and Lewy body disease have discrete pathological signatures of gait? *Alzheimers Dement.* 2019;**15**, 1367–1377.

[5] Verghese J, Robbins M, Holtzer R, et al (2008) Gait dysfunction in mild cognitive impairment syndromes. *J Am Geriatr Soc* **56**, 1244–1251.

[6] Moon Y, Sung J, An R, Hernandez ME, Sosnoff JJ (2016) Gait variability in people with neurological disorders: a systematic review and meta-analysis. *Hum Mov Sci* **47**, 197–208.

[7] Chiba H, Ebihara S, Tomita N, Sasaki H, Butler JP (2005) Differential gait kinematics between fallers and non-fallers in community-dwelling elderly people. *Geriatr Gerontol Int* **5**, 127–134.

[8] VanSwearingen JM, Paschal KA, Bonino P, Yang JF (1996) The modified gait abnormality rating scale for recognizing the risk of recurrent falls in community-dwelling elderly adults. *Phys Ther* **76**, 994–1002.

[9] Rehman RZU, Zhou Y, Del Din S, Alcock L, Hansen C, Guan Y, Hortobágyi T, Maetzler W, Rochester L, Lamoth CJ (2020) Gait analysis with wearables can accurately classify fallers from non-fallers: a step toward better management of neurological disorders. *Sensors* **20**, 6992.

[10] Craig JJ, Bruetsch AP, Huisinga JM (2019) Coordination of trunk and foot acceleration during gait is affected by walking velocity and fall history in elderly adults. *Aging Clin Exp Res* **31**, 943–950.

[11] Ritt M, Schülein S, Lubrich H, Bollheimer L, Sieber C, Gassmann KG (2017) High-technology based gait assessment in frail people: associations between spatio-temporal and three-dimensional gait characteristics with frailty status across four different frailty measures. *J Nutr Health Aging* **21**, 346–353.

[12] Kressig RW, Gregor RJ, Oliver A, Waddell D, Smith W, O'Grady M, Curns AT, Kutner M, Wolf SL (2004) Temporal and spatial features of gait in older adults transitioning to frailty. *Gait Posture* **20**, 30–35.

[13] Tura A, Raggi M, Rocchi L, Cutti AG, Chiari L (2010) Gait symmetry and regularity in transfemoral amputees assessed by trunk accelerations. *J Neuroeng Rehabil* **7**, 1–10.

[14] Fraser KC, Meltzer JA, Rudzicz F (2016) Linguistic features identify Alzheimer's disease in narrative speech. *J Alzheimers Dis* **49**, 407–422.

[15] Ahmed S, de Jager CA, Haigh AM, Garrard P (2013) Semantic processing in connected speech at a uniformly early stage of autopsy-confirmed Alzheimer's disease. *Neuropsychology* **27**, 79–85.

[16] Khodabakhsh A, Yesil F, Guner E, Demiroglu C (2015) Evaluation of linguistic and prosodic features for detection of Alzheimer's disease in Turkish conversational speech. *EURASIP J Audio Speech Music Process* **2015**,1–15.

[17] Roark B, Mitchell M, Hosom JP, Hollingshead K, Kaye J (2011) Spoken language derived measures for detecting mild cognitive impairment. *IEEE Trans Audio Speech Lang Process* **19**, 2081–2090.

[18] König A, Satt A, Sorin A, Sorin A, Hoory R, Toledo-Ronen O, Derreumaux A, Manera V, Verhey F, Aalten P, Robert PH, Renaud D (2015) Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease. *Alzheimers Dement (Amst)* **1**, 112–124.

[19] Martínez-Sánchez F, Meilán JJG, Pérez E, Carro J, Arana JM (2012) Expressive prosodic patterns in individuals with Alzheimer's disease. *Psicothema* **24**, 16–21.

[20] Meilán JJG, Martínez-Sánchez F, Carro J, López DE, Millian-Morell L, Arana JM (2014) Speech in Alzheimer's disease: Can temporal and acoustic parameters discriminate dementia? *Dement Geriatr Cogn Disord* **37**, 327–334.

[21] Boschi V, Catricala E, Consonni M, Chesi C, Moro A, Cappa SF (2017) Connected speech in neurodegenerative language disorders: a review. *Front Psychol* **8**, 269.

[22] Farrús M, Hernando J, Ejarque P (2007) Jitter and shimmer measurements for speaker recognition. in *Eighth annual conference of the international speech communication association*.

[23] Mermelstein P (1976) Distance measures for speech recognition, psychological and instrumental. *Pattern recognition and artificial intelligence* **116**, 374–388.

[24] McFee B, Raffel C, Liang D, Ellis DP, McVicar M, Battenberg E, Nieto O (2015) librosa: Audio and music signal analysis in python. in *Proceedings of the 14th python in*

science conference **8**, 18–25.

[25] Qiita. Signal-Analysis. https://brookemosby.github.io/Signal_Analysis/, Last updated Aug 24, 2017, Accessed on June 8, 2021.

[26] Honoré A (1979). Some simple measures of richness of vocabulary. *Association for literary and linguistic computing bulletin* **7**, 172–177.

[27] Croisile B, Ska B, Brabant MJ, Duchene A, Lepage Y, Aimard G, Trillet M (1996) Comparative study of oral and written picture description in patients with Alzheimer's disease. *Brain Lang* **53**, 1–19.

[28] Kavé G, Levy Y (2003) Morphology in picture descriptions provided by persons with Alzheimer's disease. *J Speech Lang Hear Res* **46**, 341–352.

[29] Janome: Japanese morphological analysis engine. https://github.com/mocobeta/janome, Last updated Sep 21, 2020, Accessed on July 14, 2021.

[30] Kawa J, Bednorz A, Stępień P, Derejczyk J, Bugdol M (2017) Spatial and dynamical handwriting analysis in mild cognitive impairment. *Comput Biol Med* **82**, 21–28.

[31] Fellows RP, Dahmen J, Cook D, Schmitter-Edgecombe M (2017) Multicomponent analysis of a digital Trail Making Test. *Clin Neuropsychol* **31**, 154–167.

[32] Werner P, Rosenblum S, Bar-On G, Heinik J, Korczyn A (2006) Handwriting process variables discriminating mild Alzheimer's disease and mild cognitive impairment. *J Gerontol B Psychol Sci Soc Sci* **61**, P228–P236.

[33] Yan JH, Rountree S, Massman P, Doody RS, Li H (2008) Alzheimer's disease and mild cognitive impairment deteriorate fine movement control. *J Psychiatr Res* **42**, 1203–1212.

[34] Müller S, Herde L, Preische O, Zeller A, Heymann P, Robens S, Elbing U, Laske C (2019) 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. *Sci rep* **9**, 1–10.

[35] Press Y, Velikiy N, Berzak A, Tandeter H, Peleg R, Freud T, Punchik B, Dwolatzky T (2012) A retrospective analysis of the sentence writing component of the mini mental state examination: cognitive and affective aspects. *Dement Geriatr Cogn Disord* **33**, 125–131.

[36] Garre-Olmo J, Faúndez-Zanuy M, Ipiña KLd, Calvó-Perxas L, Turró-Garriga O (2017) Kinematic and pressure features of handwriting and drawing: preliminary results between patients with mild cognitive impairment, Alzheimer disease and healthy controls. *Curr Alzheimer Res* **14**, 960–968.

[37] Caligiuri M, Snell C, Park S, Corey-Bloom J (2019) Handwriting movement abnormalities in symptomatic and premanifest Huntington's disease. *Mov Disord Clin Pract* **6**, 586–592.

[38] Szumilas M, Lewenstein K, Ślubowska E, Szlufik S, Koziorowski D (2020) A multimodal approach to the quantification of kinetic tremor in Parkinson's disease. *Sensors* **20**, 184.

[39] Buuren Sv, Groothuis-Oudshoorn K (2010) MICE: Multivariate imputation by chained equations in R. *J Stat Softw.* 1–68.

Supplementary Table 1: Behavioral features with statistically significant differences between diagnosis categories of AD, MCI, and CN

	CN (N=47)	MCI (N=45)	AD (N=26)	*p* Unadjusted	*p* Adjusted
Gait					
Gait speed [m/s]	1.33 ± 0.16 [A]	1.24 ± 0.21 [A]	1.07 ± 0.18 [C, M]	**<.0001**	**<.0001**
Gait speed RMS [m/s]	1.34 ± 0.17 [A]	1.25 ± 0.21 [A]	1.08 ± 0.18 [C, M]	**<.0001**	**<.0001**
Gait speed variability (SD) [m/s]	0.14 ± 0.04 [A]	0.12 ± 0.02	0.11 ± 0.02 [C]	**.0011**	**.0022**
Peak gait speed [m/s]	1.73 ± 0.35 [M, A]	1.56 ± 0.23 [C, A]	1.40 ± 0.19 [C, M]	**<.0001**	**<.0001**
Gait acceleration [m/s/s]	3.45 ± 1.50 [A]	2.96 ± 0.62 [A]	2.50 ± 0.51 [C, M]	**.0015**	**.0038**
Gait acceleration RMS [m/s/s]	3.94 ± 2.34	3.23 ± 0.66 [A]	2.80 ± 0.53 [M]	**.0089**	**.0168**
Step length [m]	0.68 ± 0.07 [M, A]	0.65 ± 0.09 [C, A]	0.58 ± 0.07 [C, M]	**<.0001**	**<.0001**
Stride length [m]	1.37 ± 0.14 [M, A]	1.30 ± 0.18 [C, A]	1.16 ± 0.15 [C, M]	**<.0001**	**<.0001**
Step time [s]	0.52 ± 0.04 [A]	0.53 ± 0.04	0.55 ± 0.05 [C]	**.0132**	**.0105**
Stride time [s]	1.04 ± 0.07 [A]	1.07 ± 0.08	1.11 ± 0.11 [C]	**.0129**	**.0098**
Stride time variability (SD) [s]	0.06 ± 0.02 [A]	0.07 ± 0.03	0.08 ± 0.03 [C]	**.0210**	**.0292**
Step time left-right asymmetry [s]	0.04 ± 0.02 [A]	0.05 ± 0.03	0.07 ± 0.05 [C]	**.0113**	**.0182**
Step angle variability (SD) [rad]	0.04 ± 0.02 [A]	0.05 ± 0.02	0.06 ± 0.03 [C]	**.0016**	**.0065**
Foot swing speed [m/s]	1.61 ± 0.17 [A]	1.53 ± 0.22 [A]	1.37 ± 0.20 [C, M]	**<.0001**	**<.0001**
Foot swing speed variability (SD) [m/s]	0.64 ± 0.05 [A]	0.61 ± 0.07	0.58 ± 0.07 [C]	**.0016**	**.0058**
Heal strike angle [rad]	0.53 ± 0.09 [M, A]	0.47 ± 0.11 [C]	0.43 ± 0.09 [C]	**.0002**	**.0007**
Toe off angle [rad]	−1.27 ± 0.12 [A]	−1.22 ± 0.13 [A]	−1.13 ± 0.11 [C, M]	**<.0001**	**.0003**
Maximum toe clearance [m]	0.13 ± 0.02 [M, A]	0.12 ± 0.03 [C]	0.11 ± 0.02 [C]	**.0010**	**.0009**

	CN (N=47)	MCI (N=45)	AD (N=26)	*p* Unadjusted	*p* Adjusted
Speech					
Honoré's statistics [$\times 10^3$] (Picture description)	1.93 ± 0.96 [A]	2.07 ± 0.98 [A]	1.34 ± 0.65 [C, M]	**.0058**	**.0061**
Total number of information units normalized by the length of the speech [words/s] (Picture description)	0.27 ± 0.10 [A]	0.26 ± 0.12 [A]	0.19 ± 0.09 [C, M]	**.0044**	**.0098**
Proportion of mistakes (Counting backwards)	0.00 (0–0.21) [A]	0.00 (0–0.27)	0.07 (0–0.80) [C]	**.0010**	**.0003**
Proportion of mistakes (Subtraction)	0.13 ± 0.17 [A]	0.20 ± 0.23	0.36 ± 0.28 [C]	**.0003**	**.0005**
Number of correct answers [words] (Phonemic verbal fluency)	7.6 ± 2.7 [A]	6.9 ± 2.7	5.6 ± 2.7 [C]	**.0194**	**.0184**
Number of correct answers [words] (Semantic verbal fluency)	17.5 ± 4.9 [A]	16.4 ± 4.6 [A]	10.6 ± 4.4 [C, M]	**<.0001**	**<.0001**
Proportion of pause duration in calculation (Counting backwards)	0.22 ± 0.08 [A]	0.25 ± 0.10	0.29 ± 0.12 [C]	**.0174**	**.0153**
Proportion of pause duration in spontaneous speech (Picture description)	0.37 ± 0.11 [A]	0.40 ± 0.10 [A]	0.50 ± 0.17 [C, M]	**.0001**	**.0003**
Proportion of pause duration in word production (Semantic verbal fluency)	0.75 ± 0.07 [A]	0.77 ± 0.05 [A]	0.84 ± 0.07 [C, M]	**<.0001**	**<.0001**
Pitch variability (SD) [Hz] (Counting backwards)	23.3 ± 9.6 [A]	19.6 ± 9.4	16.8 ± 5.4 [C]	**.0120**	**.0148**
MFCC1 (Picture description)	−496.2 ± 31.9 [A]	−492.6 ± 36.3	−514.4 ± 35.5 [C]	**.0394**	**.0273**
MFCC9 (Picture description)	−4.01 ± 6.20 [M]	−0.32 ± 6.42 [C]	−2.16 ± 5.01	**.0183**	**.0152**
MFCC9 skewness (Picture description)	−0.06 ± 0.16 [M]	0.04 ± 0.18 [C]	−0.01 ± 0.19	**.0326**	**.0189**

	CN (N=47)	MCI (N=45)	AD (N=26)	p Unadjusted	p Adjusted
Drawing					
Drawing speed [mm/s] (TMT-B)	102.0 ± 18.5 A	100.5 ± 28.4 A	79.0 ± 20.8 C, M	**.0002**	**.0006**
Drawing speed variability (SD) [mm/s] (TMT-B)	34.8 ± 7.6 A	34.9 ± 13.6 A	25.3 ± 12.7 C, M	**.0015**	**.0042**
Maximum drawing speed [mm/s] (TMT-B)	135.9 ± 26.3 A	130.5 ± 41.1 A	102.0 ± 33.8 C, M	**.0004**	**.0013**
Total stroke length [mm] [$\times 10^2$] (Sentence)	8.46 ± 3.89 A	9.50 ± 5.38	6.17 ± 2.61 C	**.0100**	**.0215**
Pause duration between drawings [s] (TMT-A)	0.60 ± 0.27 A	0.74 ± 0.30 A	1.09 ± 0.65 C, M	**<.0001**	**<.0001**
Pause duration between drawings [s] (TMT-B)	1.49 ± 0.56 M, A	1.98 ± 0.84 C, A	2.55 ± 0.98 C, M	**<.0001**	**<.0001**
Maximum pause duration between drawings [s] (TMT-A)	3.83 ± 2.98 A	5.25 ± 4.12 A	8.30 ± 5.79 C, M	**.0002**	**.0007**
Maximum pause duration between drawings [s] (TMT-B)	7.61 ± 4.28 M, A	15.72 ± 15.88 C	20.41 ± 14.61 C	**<.0001**	**.0002**
Pause duration between drawings normalized by total stroke length [s/mm] (CDT)	0.023 ± 0.009 A	0.031 ± 0.030	0.045 ± 0.039 C	**.0040**	**.0121**
Pause duration between drawings normalized by total stroke length [s/mm] (Pentagon)	0.027 ± 0.015 A	0.032 ± 0.021	0.043 ± 0.029 C	**.0104**	**.0328**
Pause duration between drawings normalized by total stroke length [s/mm] (Sentence)	0.016 ± 0.009 A	0.019 ± 0.012	0.025 ± 0.013 C	**.0057**	**.0154**
Pause duration between drawings normalized by total stroke length [s/mm] (TMT-A)	0.011 ± 0.007 A	0.016 ± 0.010 A	0.025 ± 0.019 C, M	**<.0001**	**.0001**
Pause duration between drawings normalized by total stroke length [s/mm] (TMT-B)	0.030 ± 0.013 M, A	0.052 ± 0.038 C, A	0.088 ± 0.047 C, M	**<.0001**	**<.0001**

	CN (N=47)	MCI (N=45)	AD (N=26)	p Unadjusted	p Adjusted
Task duration normalized by total stroke length [s/mm] (CDT)	0.036 ± 0.010 ^{A}	0.045 ± 0.032	0.059 ± 0.039 ^{C}	**.0031**	**.0097**
Task duration normalized by total stroke length [s/mm] (Sentence)	0.033 ± 0.013 ^{A}	0.035 ± 0.015	0.045 ± 0.017 ^{C}	**.0055**	**.0160**
Task duration normalized by total stroke length [s/mm] (TMT-A)	0.021 ± 0.007 ^{A}	0.026 ± 0.011 ^{A}	0.037 ± 0.021 $^{C, M}$	**<.0001**	**<.0001**
Task duration normalized by total stroke length [s/mm] (TMT-B)	0.040 ± 0.014 $^{M, A}$	0.063 ± 0.039 $^{C, A}$	0.102 ± 0.047 $^{C, M}$	**<.0001**	**<.0001**
Pressure variability (CV) (TMT-A)	0.060 ± 0.027 ^{A}	0.093 ± 0.064	0.097 ± 0.063 ^{C}	**.0029**	**.0249**
Pressure variability (CV) (TMT-B)	0.077 ± 0.037 $^{M, A}$	0.124 ± 0.068 ^{C}	0.149 ± 0.074 ^{C}	**<.0001**	**<.0001**
Number of correct edges (TMT-B)	23.4 ± 1.2 $^{M, A}$	21.7 ± 3.9 $^{C, A}$	14.2 ± 10.0 $^{C, M}$	**<.0001**	**<.0001**
Number of incorrect edges (TMT-B)	0.7 ± 1.1 $^{M, A}$	2.0 ± 2.9 ^{C}	4.3 ± 5.0 ^{C}	**<.0001**	**.0001**
Number of lifts (TMT-A)	3.3 ± 4.0 ^{A}	7.7 ± 8.7	8.2 ± 7.4 ^{C}	**.0032**	**.0243**
Number of lifts (TMT-B)	6.6 ± 5.7 $^{M, A}$	16.5 ± 17.4 ^{C}	19.8 ± 17.8 ^{C}	**.0002**	**.0017**
Time duration between nodes [s] (TMT-A)	1.10 ± 0.22 ^{A}	1.20 ± 0.36	1.41 ± 0.50 ^{C}	**.0028**	**.0107**
Time duration between nodes [s] (TMT-B)	2.36 ± 0.61 ^{A}	2.13 ± 0.87 ^{A}	2.94 ± 1.63 $^{C, M}$	**.0073**	**.0053**
Task duration normalized by the number of answered edges [s] (TMT-A)	1.52 ± 0.57 $^{M, A}$	2.06 ± 0.97 $^{C, A}$	2.90 ± 1.81 $^{C, M}$	**<.0001**	**<.0001**
Task duration normalized by the number of answered edges [s] (TMT-B)	3.59 ± 1.23 $^{M, A}$	6.09 ± 4.23 $^{C, A}$	10.28 ± 5.27 $^{C, M}$	**<.0001**	**<.0001**

Normally distributed data are displayed as mean ± standard deviations. Data for transformed variables are displayed as median (minimum–maximum) and refer to the non-transformed values. Bold values highlight statistically significant differences between CN, MCI, and AD in the unadjusted model and the adjusted model controlling for age and sex. Pairwise multiple comparisons (Bonferroni adjusted *p* values) were performed when comparing individual diagnostic groups. Significant differences between diagnosis categories are marked with C, M, or A (C: Different to CN, M: Different to MCI, A: Different to AD).

SD, Standard deviation; CV, Coefficient of variation; RMS, Root mean square; TMT-A, Trail making test-part A; TMT-B, Trail making test-part B; CDT, Clock Drawing Test, Sentence and Pentagon: Writing a sentence about anything and the copy intersecting-pentagon item of MMSE.