Cognitive Computation (2019) 11:576–586
https://doi.org/10.1007/s12559-019-09642-2

A Handwriting-Based Protocol for Assessing Neurodegenerative Dementia

Donato Impedovo[1] · Giuseppe Pirlo[1] · Gennaro Vessio[1] · Maria Teresa Angelillo[2]

Received: 24 January 2018 / Accepted: 12 April 2019 / Published online: 2 May 2019

Abstract

Handwriting dynamics is relevant to discriminate people affected by neurodegenerative dementia from healthy subjects. This can be possible by administering simple and easy-to-perform handwriting/drawing tasks on digitizing tablets provided with electronic pens. Encouraging results have been recently obtained; however, the research community still lacks an acquisition protocol aimed at (i) collecting different traits useful for research purposes and (ii) supporting neurologists in their daily activities. This work proposes a handwriting-based protocol that integrates handwriting/drawing tasks and a digitized version of standard cognitive and functional tests already accepted, tested, and used by the neurological community. The protocol takes the form of a modular framework which facilitates the modification, deletion, and incorporation of new tasks in accordance with specific requirements. A preliminary evaluation of the protocol has been carried out to assess its usability. Successively, the protocol has been administered to more than 100 elderly MCI and match controlled subjects. The proposed protocol intends to provide a "cognitive model" for evaluating the relationship between cognitive functions and handwriting processes in healthy subjects as well as in cognitively impaired patients. The long-term goal of this research is the development of an easy-to-use and non-invasive methodology for detecting and monitoring neurodegenerative dementia during screening and follow-up.

Keywords Handwriting behavior · Neurodegenerative disease · Computer-aided diagnosis · Handwriting analysis · Acquisition protocol

Introduction

Dementia is a general term for referring to a wide range of symptoms, associated with memory and thinking skills, severe enough to interfere with everyday life activities [1]. Alzheimer's disease (AD) is the most common type of dementia, accounting for 60 to 80% of cases [2]. Parkinson's disease (PD) [3] is linked to dementia too: although it primarily results in motor deficits, it also affects attention, memory, and cognitive skills [4]. Neurodegenerative disorders affect the structure and functions of brain regions resulting in a progressive cognitive, functional, and behavioral decline. Unfortunately, there is no cure and the decline can only be somehow managed during the disease progression. However, an early diagnosis is crucial in the perspective of proper medical treatment and can improve the quality of patient life. Moreover, the assessment of signs of a specific disease is useful for its diagnostic differentiation with respect to similar disorders and for monitoring its progression. To this end, special attention is devoted to mild cognitive impairment (MCI), since in this case, there is a high risk of developing dementia [5].

The evaluation of the patient's clinical status and his/her responsiveness to medication is typically achieved via neurological assessment, including various tests and rating scales. Mini-Mental State Examination (MMSE) is used extensively to assess cognitive impairment [6]. A growing research interest has arisen, in the last years, towards the application of biometric techniques to health. Traditionally, biometrics has been successfully applied to security. However, established techniques, which use biometric information within security-related applications, can be profitably used in the health domain [7]. For example, the potential of applying speech and emotional state recognition to the AD diagnosis has been recently explored [8].

✉ Donato Impedovo
donato.impedovo@uniba.it

[1] Dipartimento di Informatica, Universita degli Studi di Bari, Via E. Orabona, 4, 70125 Bari, Italy

[2] Unit in Affective and Social Neuroscience, Department of Psychology, Catholic University of Milan, Largo Gemelli, 1, 20123 Milan, Italy

As part of neuropsychological test batteries, a special role in the context of dementia assessment can be assumed by handwriting. Handwriting, in fact, is a complex activity entailing cognitive, kinesthetic, and perceptual-motor components [9], whose changes can be a prominent biomarker for the evaluation of degenerative dementia [10]. Some works (e.g., [11–13]) provided evidence that the automatic discrimination between unhealthy and healthy people can be accomplished based on kinematic features obtained through simple and easy-to-perform handwriting tasks. Handwriting has been related to AD [12, 14, 15] as well as to PD [16–22]. Depending on the complexity of the task to be carried out, handwriting involves motor aspects (e.g., slowness and rigidity of movement) as well as cognitive effort (e.g., attention and memory); therefore, it provides a way to evaluate impairments for both aspects. Developing a handwriting-based decision support system is desirable, as it can provide a complementary approach to the pathology evaluation performed by expert neurologists that is quantitative, non-invasive, and very low cost [10].

Commercially available tools (tablets) can be used to acquire data. Raw data depends on the specific device; however, the most common are (*x*–*y*) coordinates, time stamps, pen orientation (azimuth and altitude), pressure, and button status. The button status is used to identify in-air movements (when the tip is not in contact with the writing surface) and on-pad movements. It has been demonstrated that in-air movements are related to functional decline, as well as to difficulties in activity planning [12, 23–25]. Handwriting has been successfully employed to identify malingering in health care, i.e., the false information given by patients about their health [26]. Preliminary results suggested that a computerized tool based on handwriting can help detect deception. A similar tool has been recently used for capturing cognitive load implications during complex figure drawing [27, 28].

However, although relevant results have been recently obtained, several open issues demand further research [10, 29]. Many researchers have adopted databases/datasets built by collecting data themselves. First, these datasets are different in writing/drawing tasks adopted, and some controversial results are due to the fact that some evidence of impairments are connected with the same specific task and less or not with others [10, 29]. Second, available datasets are very reduced in size (the number of participants). Third, the elicitation of the diagnosis is never mentioned and the set of available metadata (e.g., age and schooling factors) is very reduced. Fourth, a single acquisition session is generally available. These lacks significantly limit research development. To overcome these problems and to be able to collect a big dataset, specific attention must be paid to the acquisition protocol. To address this issue, this paper aims to propose a handwriting-based protocol for neurodegenerative dementia assessment. At the same time, since this proposal intends to be the first step towards the development of an integrated computer-aided system, it includes also a digitized version of standard tests already accepted, tested and used by the neurological community working in this field that will be used to have (in the initial stage of the development) the ground truth. Note that these tests already involve handwriting tasks: they can be subjected to kinematic analysis to provide additional data to those that can be obtained by their pen- and paper-based counterpart. On the other hand, the proposal includes handwriting experimental tasks, useful for extracting kinematic features, which are currently under investigation by researchers.

Proposed Protocol

A typical clinical workup includes some core elements: the medical history documents family history, and previous illnesses. The assessment of independent function and daily activities allows the examiner to focus on changes in the usual level of the individual. Neuropsychological testing is also carried out, so that specific thinking skills can be evaluated through series of written tests. The protocol here proposed is the result of previous research:

- an in-depth study of the state of the art of pattern recognition approaches on handwriting/drawing tasks related to this field [10];
- an in-depth study on the relation between tasks and disease evidence [29];
- an overview and understanding of already accepted standard tests for the assessment of neurodegenerative diseases [30].

Therefore, this work is a synthesis of previous works in which a specific protocol is proposed, implemented, and tested given a set of clinical and research evidence.

In light of the observations above, the proposed protocol is made up of three parts:

- an initial screening,
- a battery of a digitized version of standard cognitive and functional tests;
- some of the most relevant handwriting tasks extracted from the research literature.

The first two parts mostly require the examiner to provide a data entry; the last one requires the subject to perform the tasks. An overall model of the protocol is depicted in Fig. 1. The following subsections describe the three main components of the proposal.

Initial Screening

The first part of the protocol consists of a simple screening section allowing the examiner to collect personal and clinical information about the subject. This part serves also to match

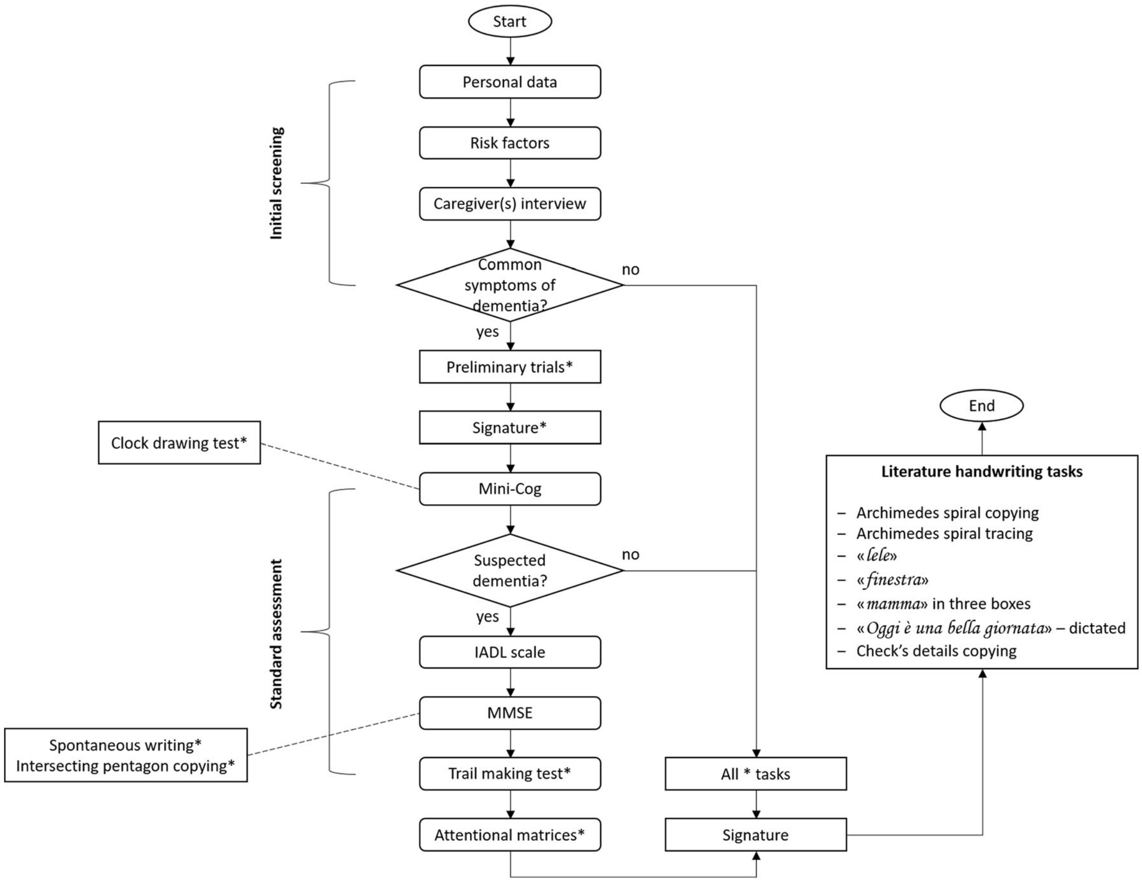

Fig. 1 Overall model of the proposed protocol. The use of the stars emphasizes how the handwriting tasks have to be administered not only to impaired patients but also to healthy controls in order to provide examples for both categories

data from the same subjects in separate sessions: in this way, the progression of the disease and the cycle of medication to which patients are subjected can be monitored.

Personal Data The following data are acquired: id; gender; level of education (expressed in years); current or past employment (intellective or manual); drugs taken, and family history. These data shall be confidential; moreover, their treatment, according to the current legislation, must be authorized. In addition, the name and surname of the examiner and the place and date of the test are recorded.

Risk Factors The presence of risk factors is evaluated. They include diabetes, hypertension, obesity, smoke, depression, physical inactivity.

Caregiver(s) Interview Caregivers are asked to answer questions about the most common symptoms of dementia. More specifically, the followings can be considered:

- Memory loss;
- Difficulty in performing familiar tasks;
- Problems with language;
- Disorientation to time and place;
- Poor or decreased judgment;
- Problems with keeping track of things;
- Misplacing things;
- Changes in mood or behavior;
- Changes in personality;
- Loss of initiative.

If at least one question of the interview receives a positive answer, the protocol continues; otherwise, it ends.

Preliminary Trials It is worth noting that the use of a digitizing tablet may be unfamiliar to elderly people. To mitigate this problem, the protocol requires the subject to perform some simple tasks. These trials are intended to provide a simplified version of some of the tasks the subject is going to perform. They include:

- Writing the word ciao ("hello" in Italian);
- Connecting two horizontal points with a straight
- line repeatedly (for four times);
- Connecting two vertical points with a straight line repeatedly (for four times);
- Copying a square.

Signature (First) The subject is asked to provide her/his signature. The signature is acquired both in this initial stage and at the end of the second stage, as signatures may vary depending on the physical and emotional state of the signer [31], especially as an effect of fatigue.

Standard Assessment

A standard assessment typically includes cognitive and functional tests. Tests included in the protocol are Mini-Cog [32]; the Instrumental Activities of Daily Living (IADL) scale [33]; the already mentioned MMSE; the trail making test [34], and the attentional matrices test [35]. These tests have been included in the proposed protocol since they are among the most widely used and acknowledged in the clinical settings for evaluating dementia and cognitive impairment.

It is worth noting that digitizing such tests provides benefits to neurologists under several viewpoints, in particular paper saving and data storage (data can be accessed easily in a structured form and possibly from anywhere). These tests are used to provide the ground truth diagnosis of acquired data.

Mini-Cog It is a short screening test which is composed of two parts: a 3-item recall test and a clock drawing test. To verify short-term memory loss, the test evaluates visuospatial abilities, attention, and executive functions. The score is calculated as follows: 1 point for each word recalled without cues and 2 points for a normal clock drawing. A total score of 3, 4, or 5 indicates a lower likelihood of dementia. Mini-Cog is not a diagnostic test, but it is useful to identify subjects at risk of developing dementia and for which a more in-depth assessment is needed. Therefore, one may proceed with the administration of the following tests only after obtaining a low Mini-Cog score.

IADL Scale This is an instrument to assess independent living skills. More precisely, it provides information about eight fundamental skills such as, for example, food preparation, housekeeping, and laundering. Individuals are scored in accordance with their highest level of functioning and on the need for assistance or supervision. The evaluation by the examiner is based on information provided by the subject themselves, if they are cognitively capable, or by the caregiver. The scale assigns a score for each function: 0 (dependent) or 1 (independent). A summary score, therefore, ranges from 0 (low function) to 8 (high function).

MMSE It consists of a 30-point questionnaire including questions and problems in many areas: orientation to time and place; repetition of lists of words; and attention and calculation. MMSE is one of the most widely used tests for both dementia diagnosis and the evaluation of its degree of severity. As for the hand-drawing/handwriting task, the test asks to write a spontaneous sentence and copying two intersecting pentagons. Any score greater than or equal to 24 points indicates a normal cognition. Below this, the score can indicate severe (≤ 9 points), moderate (10–18 points), or mild (19–23 points) cognitive impairment. Note that the MMSE score is automatically corrected in accordance with the age and the level of education previously collected during the initial assessment.

Trail Making Test It requires the subject to connect a sequence of 25 consecutive targets. The test is divided into two parts: in the first part, the targets are only numbers (1, 2, 3, …); in the second part, the subject alternates between numbers and letters (1, A, 2, B, …). Note that two trails are provided before the effective task is evaluated. The test is tailored to explore different cognitive components: attentional skills, visuomotor planning, and working memory [36].

Attentional Matrices It is a cancelation test: three matrices, i.e., grid of numbers, are shown to the test taker and he/she is asked to cross out as fast as possible target numbers of either one, two, or three digits. The aim of the test is to evaluate the subject's ability to detect visual targets among several distractors.

The abovementioned tests already include handwriting tasks, and their importance within the tests has been already clinically validated. In this case, the use of a digital acquisition tool gives the possibility to acquire kinematic features, providing more information than their traditional pen-and-paper counterpart. Moreover, it is worth remarking that the inclusion of standardized tests helps in specifying the stage of disease of the patterns acquired, resulting in a multiclass classification problem. In fact, they provide a "cognitive model" which enables the evaluation of the relationship between cognitive functions and handwriting processes in healthy and unhealthy subjects. If and to which extent changes in the handwriting process could be considered as a potential biomarker of ongoing cognitive impairments can also be evaluated.

Signature (Second) In conclusion of the second stage, a second signature of the subject is acquired.

Research Drawing/Handwriting Tasks

The scientific community has inspected many different drawing/writing tasks can be classified according to the following taxonomy [10]:

- Simple drawing tasks. The interest is into performing simple measures of trajectory, size, velocity, acceleration, and tremor. Drawings such as straight lines, spirals, circles, and squares have been adopted.
- Simple writing tasks. In this case simple non-sense words, words, and short sentences have been considered. A very popular pattern is the repetition of "le" [37]. It is worth

noting that "e" and "l" both contain an up- and a down-velocity stroke but scaled in size. Moreover, degradation of handwriting has been observed in repeated actions (see Fig. 2).

In general, words and sentences are selected according to their simple orthography and easy syntax. A sentence requires a high degree of simultaneous processing and may have a higher neuro-motor programming load than a sequence of identical cuttings. It also offers the possibility to evaluate in-air movements between words as well as fatigue during writing. Handwritten signatures have been considered too [38].

- Complex tasks. These tasks involve various neuropsychological functions: auditory perception, auditory memory, abstraction capacity, visual memory, visual perception, visual-space functions, programming, and execution capacity. Some examples are copying the details of a bank check into the appropriate places and the Clock Drawing Test (CDT).

Among the various tasks already mentioned, the following were considered based on their performance and included in the proposed protocol.

Archimedes Spiral Copying The Archimedes spiral is one of the most popular handwriting exercises. Spiral drawing, in fact, has been frequently used for the evaluation of the motor performance in various movement disorders, especially PD [39, 40]. The work of Pullman pioneered the application of spiral analysis to the measurement of tremor [41]. In this specific task, an Archimedes spiral is shown to the test taker and the task simply consists in copying it.

Archimedes Spiral Tracing Differently from the previous task, the subject is required to trace the contours of the spiral.

Repetition of the Cursive Letters "le" Patterns of the cursive letters "le".

Writing in Cursive the Word "Finestra" ("Window") The word "finestra" includes ascending and descending strokes, more it can be written without lifting the pen from the surface.

Writing in Cursive, in Three Boxes, the Word "mamma" ("mother" in Italian) This is one of the first learned and remains one of the last word used before dying. It was used in [42], but the task here is slightly different: the user is asked to write consecutively the same word within three rectangular boxes of increasing size. The height of the boxes is 1, 2, and 4 cm, respectively. The aim is to evaluate the ability to scale the handwritten trace depending on the visual feedback provided. This task can be useful to evaluate fine motor control.

Writing in Cursive the Dictated Sentence "Oggi è una bella giornata" ("It is a beautiful day") This sentence has been chosen because it does not provide linguistic or spelling difficulties. Note that, to keep the dictation speed constant, an audio trace is used instead of the examiner's voice.

Check's Details Copying This is a functional task related to daily activities and involves cognitive effort [43],

Clock Drawing Test The user is required to draw a clock with numbers and dials. It has an important role in the early screening of cognitive impairment, especially in dementia [13, 44].

Implementation

The system displays a set of windows, each conceived to acquire data from either the examiner or the subject performing the handwriting/hand-drawing task. In this way, the system takes the form of a modular framework which facilitates the modification, deletion, and incorporation of new tasks, in accordance with specific needs and requirements, by simply modifying, deleting, or adding windows. The system not only interacts with the users but also interacts with an underlying database which records metadata. The following subsections describe the main implementation issues.

Device

The protocol has been developed on the commercially available Wacom MobileStudio Pro 13, which is a full-featured tablet with visual feedback and computational capabilities (the main features of the tablet are summarized in Table 1). Contrary to other hardware, therefore, it immediately provides visual feedback and does not require to be connected to a

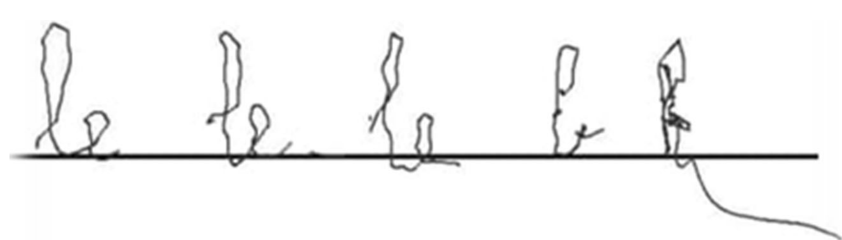

Fig. 2 Repetitions of "le"

Table 1 Main features of Wacom MobileStudio Pro 13

Feature	Specification
Size	367 × 229 × 16 mm
Weight	1420 g
Display	13.3 in.
Displayable colors	16.7 million
Technology	LED
Resolution	WQHD (2560 × 1440)
Operating system	Windows 10 Pro
Processor	Intel Core i7
Graphics card	Intel Iris Graphics 550
Storage	256GB (SSD)
RAM	8GB DD3
Pen	Wacom Pro Pen 2
Multi-touch	Yes
Pressure levels	8192 levels
Tilt range	60°
Tilt recognition	± 60 levels
Sampling frequency	200 Hz

desktop computer: this makes data acquisition more natural and easier. The protocol program has been coded in Java, version 8 update 144, within the integrated development environment Eclipse Oxygen. Among the others, the program makes use of the JPen software library, which allows capturing of data samples through the low-level Wintab API. Data acquired by the protocol are stored in .txt files that can be easily accessed by querying a relational database built with MySql version 5.7.20.

Acquisition

The program is conceived for two acquisition scenarios:

- the overall protocol is administered so that a first diagnosis of the subject can be carried out;
- the administration of only the graphical tasks (i.e., those requiring the interaction between the test taker and the pen), so that the data of follow-up study visits can be collected.

For the latter, the administration of the standard tests is assumed to have already been conducted. Moreover, it is worth noting that administering the overall protocol may require a long time; to alleviate the effect of fatigue, the program allows the examiner to skip some tasks which can be resumed later or even in another day.

The raw data that can be captured by the device are the *x*- and *y*-coordinates of the pen position and their timestamps, pen inclination (tilt-*x* and tilt-*y*), and pen pressure. Finally, the tablet is also the so-called button status, which is a binary variable evaluating 0 for pen-up state (in-air movement) and 1 for pen-down state (on-surface movement). All the features have the same length, varying from execution to execution. Therefore, the execution of a task can be described by a matrix $X = (x, y, p, t, \text{tilt\_}x, \text{tilt\_}y, b)$, where each column is a vector of length N, while N equals to the number of sampled points.

Figure 3 depicts the acquisition process. The device captures more information than the handwritten trace left on the surface (a, b), which is stored on a separate text file (c).

Human-Computer Interface

The interface has been designed for a face-to-face interaction between the examiner and the "patient." In the typical acquisition scenario, the examiner is sitting on one side of the table, entering data, making questions, explaining exercises, and, more in general, conducting the test. The test taker is sitting on the opposite side of the table and the tablet. Thanks to this multimodal interface, the system provides the users with a combination of input and output modalities to communicate with it. This makes the application intuitive and handy not only for expert neurologists but also for informal caregivers as well as the caregivers working in day centers for aging people.

All drawing tasks (e.g., spiral copying) provide an active area of ~ 28.5 × 12.5 cm. All the writing tasks provide an active area that resembles a traditional A4 paper, so it is ~ 21 × 11 cm. The signature tasks are characterized by a smaller input area, which corresponds to the display size (~ 9.6 × 6 cm) of the Wacom STU-430, which is a device specifically tailored to acquire signatures for future pattern matching. Note that the program is conceived for sampling data even when the subject writes outside the acquisition area margins; however, in such cases, no visual feedback is provided. For what concerns trail making test and attentional matrices, a stopwatch is displayed on the window to keep track of the time of completion.

Experimental Results

Preliminary Evaluation

A usability evaluation has been performed before starting real data acquisition. More precisely two methodologies were adopted:

- A System Usability Scale (SUS) test [45] was administered to 10 healthy young adults (average age 23);
- A thinking aloud protocol [46] was conducted with 10 elderly people (average age 70). It is worth remarking that

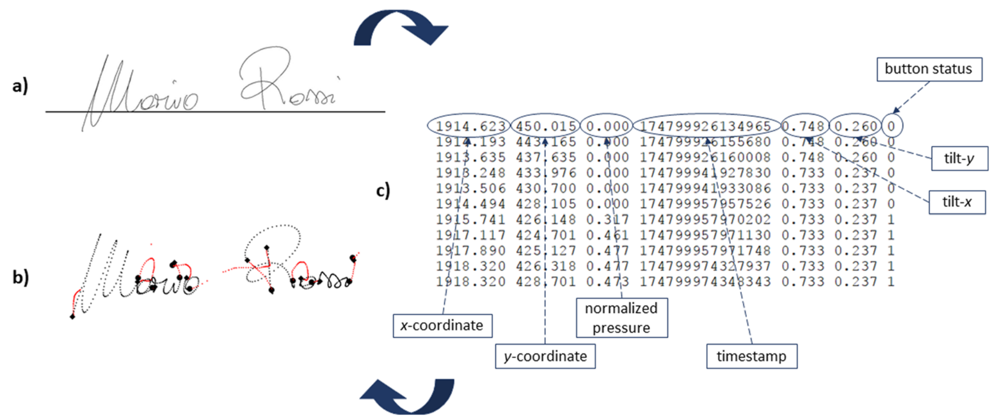

Fig. 3 **a** The handwritten trace of a fake signature written by a healthy young adult left on the tablet surface. **b** A small portion of the corresponding raw data sampled by the tablet. **c** The reconstruction of (**a**) from (**b**), including in-air movements. **c** Red dots represent in-air movements, diamonds represent pen-downs, circles indicate pen-ups

the latter did not undergo a neuropsychological evaluation to assess dementia, but they were supposed to be healthy.

For both tests, the users were requested to perform only the handwriting/hand-drawing tasks. No training process was scheduled before the evaluation and users started the tests without previous knowledge. The only way they had to become familiar with the system was the preliminary trials already included in the protocol. This part of the test had the aim to evaluate activity understanding level by a healthy cohort.

The SUS test is a well-known usability test that is very easy to administer. It consists of a 10-item questionnaire with 5 possible answers for each question: they range from "strongly disagree" to "strongly agree." Table 2 shows the results of the evaluation carried out with the young adult sample. The test yields a single score on a scale of 0–100 (not to be intended as a percentage) calculated as follows: for each of the odd-numbered questions, subtract 1 from the score; for each of the even-numbered questions, subtract their value from 5; add up the total score; multiply this score by 2.5. The global score obtained averaging all over the 10 users was 84.8, which was quite high considering that a score above 68 is considered above average.

The SUS test is less tailored to elderly people, as they have less familiarity with technological issues. For this reason, for only the elderly sample, a thinking aloud-based evaluation was adopted. It simply requires the subjects to verbalize their thoughts as the task proceeds. Users reported they felt comfortable using the electronic pen for writing; moreover, they did not feel any difference with respect to the traditional pen- and paper-based writing.

Field Evaluation

After the initial usability evaluation and some refining, the tool has been used to acquire data on the field. Up to date, 105 participants (67 females and 38 males) have been involved. Table 3 reports details about this new set of users.

Table 4 shows the number of users has been able to complete the test skipping 0, 1, etc. number of handwriting/drawing tasks. The bar chart of Fig. 4 and Table 5 report tasks time execution details.

According to the proposed protocol, the dementia assessment revealed that 71 users out of 105 were cognitively impaired. Table 6 reports age details for the two groups.

From the time series raw data sampled by the acquisition device, the following measures, i.e., features, have been extracted:

- The number of on-surface strokes: an on-surface stroke corresponds to the consecutive points sampled between a pen-down and the following pen-up;
- On-surface time: overall time spent on-surface;
- In-air time: overall time spent in-air;
- Total time: overall time for task completion;
- Mean pressure: average pressure over all on-surface strokes.

These features have been used to classify cognitively impaired vs healthy subjects. For each user, the set of all features for each task have been lumped into one multidimensional vector. The Random Forest classifier has been used with a 5-fold cross validation. Since the dataset is (at this stage) unbalanced, results have been evaluated in terms of precision, recall, and *F*-measure. Precision is intuitively the ability of the classifier to not label as positive a sample that is in the negative class. Recall is the ability of the model to find all the positive examples. *F*-measures can be interpreted as a weighted harmonic mean of precision and recall. Results obtained are reported in Table 7.

Discussion

Handwriting/drawing tasks offer the possibility to investigate neurodegenerative diseases. The proposed protocol:

Table 2 Results of the SUS test. Answers are encoded in such a way that 1 stands for "strongly disagree" and 5 for "strongly agree"

Question	Avg. value	Best value	Worst value
Q1: "I think that I would like to use this system frequently"	3.4	5	1
Q2: "I found the system unnecessarily complex"	1.2	1	2
Q3: "I thought the system was easy to use"	4.4	5	3
Q4: "I think that I would need the support of a technical person to be able to use the system"	1.8	1	4
Q5: "I found the various functions in this system were well integrated"	4.5	5	4
Q6: "I thought there was too much inconsistency in this system"	1.4	1	3
Q7: "I would imagine that most people would learn to use this system very quickly"	4.6	5	4
Q8: "I found the system very cumbersome to use"	1.5	1	2
Q9: "I felt very confident using the system"	4.2	5	3
Q10: "I needed to learn a lot of things before I could get going with this system"	1.3	1	2

- Gives the possibility to doctors to formulate a dementia diagnosis according to a set of well-known assessments eventually administered in multiple sessions;
- Gives the possibility to acquire handwriting dynamics related to multiple motor, functional and cognitive experimental tasks.
- According to the second point of the previous list, the viability of the proposed protocol has been demonstrated by the fact that:

– 91.43% of users were able to complete the test skipping less than 3 tasks;
– The set of tasks required a mean time of 5.74 (m) with a std. of 2.82 (m).

The overall min execution time has been 2 (min), the max 21 (min) with a mean of 5.74 (min).

The time duration analysis of single task reveals that bank-check copying, the second trail making test and the third attention matrix were the heaviest load tasks, in fact they are complex tasks embedding motor, cognitive, functional and execution aspects. However, the bank-check copying task and the third attention matrix were completely performed, respectively, in 86.7% and 94.3% of cases, while more than half of patients were unable to perform the second trail-making test. This result calls for a deep investigation since these tasks could be useful to identify the stage of the disease as well as its evolution. Table 3 also reveals that signature time increases between the first and the second trial thus unveiling the fatigue. Concerning the clock drawing test, it has been observed that people with better cognitive-functional level generally divide the circle into different quadrants, placing the numbers 12–3–6–9 and then the others. Conversely, patients with dementia start writing from 1 or 12 and they proceed to fill the whole space with other numbers; often, the clock is filled leaving out either the first or the last number, so that even if time execution seems to be not discriminant, the execution itself is very poor or completely wrong. Also, these results suggest the development of specific approaches in order to investigate the specific task. Finally, regarding the trail-making test, the examiner is interested in evaluating the time of completion and the number of errors; however, handwriting dynamics are also able to unveil the search pattern adopted by the user.

The classification of healthy and impaired users exhibited high recall for the HC class and low recall for the MCI class: this indicates that such an approach may be better in identifying the absence of illness in the healthy population rather than the presence of illness in the pathological group. However, these are preliminary results which call for further processing and deeper analysis.

Table 3 Users details

Age			
Min	Max	Mean	Std
45	99	75.8	14.6

Table 4 The number of users and skipped tasks

No. of users	No. of skipped tasks
0	43
1	19
2	26
3	8
4	0
5	3
6	1
7	1
8	2
9	0
10	1
11	0
12	0
13	1

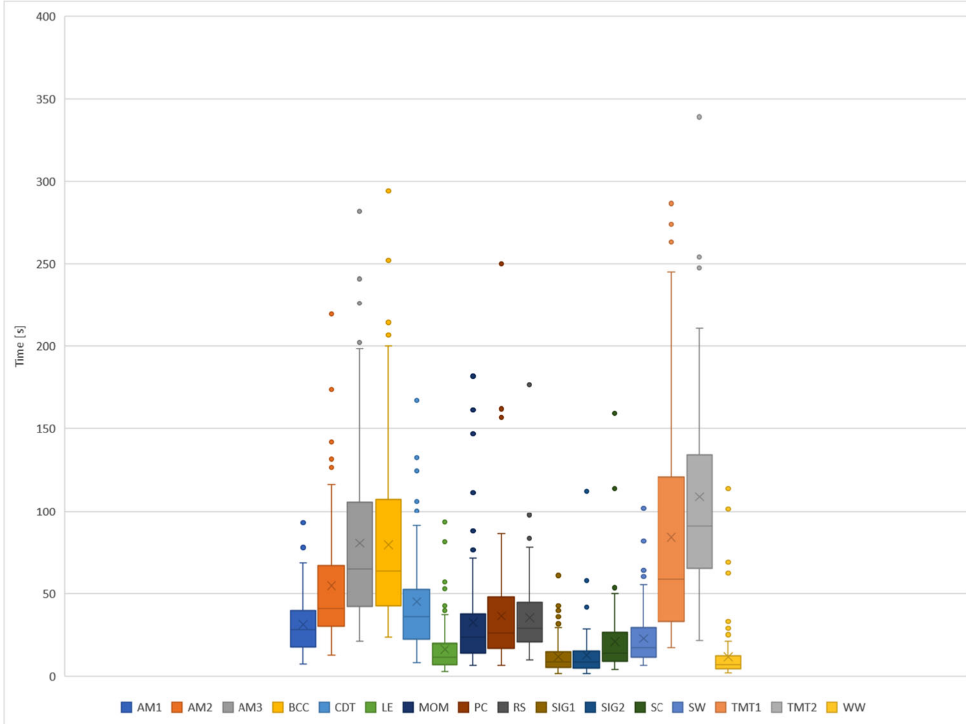

Fig. 4 Task execution time details. AM1—first attention matrix; AM2—second attention matrix; AM3—third attention matrix; BCC—bank check copying; CDT—clock drawing test; LE—"le" repetitions; MOM—"mamma"; PC—pentagons copying; RS—retraced spiral; SIG1—first signature acquisition; SIG2—second signature acquisition; SC—spiral copying; SW—"Oggi è una bella giornata"; TMT1—first trail-making test; TMT2—second trail-making test; WW—"finestra"

Conclusion and Future Work

This paper proposes a protocol tailored to the assessment of neurodegenerative dementia (and mild cognitive impairment as well). It integrates handwriting tasks to a digitized version of standard cognitive and functional tests already accepted, tested, and used by neurologists. On one hand, the protocol may be useful, to the research community, to collect handwritten traits and the diagnosis ground truth helpful to design computer-aided detection systems based on handwriting. At the same time, the protocol may be of real use for neurologists to support their daily activities. The inclusion of standardized tests is important to determine the stage of the illness of the patterns acquired: indeed, the main open problem is to support the early diagnosis of the disease, as well as the monitoring of the development of the disease as it advances. To date, more than 100 patients have already completed the protocol and acquisition sessions are going and will involve those already acquired to monitor the evolution of the disease. The next step is to perform data acquisition along with a comprehensive evaluation of recognition performance also considering an extended set of features and classifiers [10]. To the best of our knowledge, this is one of the first attempts to provide an integrated protocol for data acquisition in this specific field.

Tablet technology enables the implementation of a multimodal interaction system in which not only the input provided by the electronic pen but also tactile, speech, or visual input can be acquired. Indeed, some tasks that require the interaction between the finger and the touch screen of the tablet are also under investigation. Recent literature reported on promising results using analogous tasks for PD assessment [47]. Note that such characteristics make the protocol well suited also to smartphones. Moreover, thanks to such a multimodal interface, the development of a mobile conversational agent appears to be feasible. An

Table 5 Tasks execution stats (s). The last row reports the percentage of completely executed trials. *AM1*—first attention matrix; *AM2*—second attention matrix; AM3—third attention matrix; *BCC*—bank check copying; *CDT*—clock drawing test; *LE*—"le" repetitions; *MOM*—"mamma"; *PC*—pentagons copying; *RS*—retraced spiral; *SIG1*—first signature acquisition; *SIG2*—second signature acquisition; *SC*—spiral Copying; *SW*—"Oggi è una bella giornata"; *TMT1*—first trail-making test; *TMT2*—second trail-making test; *WW*—"finestra"

	AM1	AM2	AM3	BCC	CDT	LE	MOM	PC	RS	SIG1	SIG2	SC	SW	TMT1	TMT2	WW
Min	7.6	12.6	20.8	23.3	8.1	2.8	6.7	6.5	9.8	1.8	1.7	3.9	6.6	17.3	21.6	2.08
Max	92.7	219.5	281.7	504.8	169.5	93.2	181.6	250.1	176.5	60.8	112.1	159.2	101.8	286.4	496.6	113.9
Avg	31.1	54.7	80.3	84.0	44.9	16.4	32.3	36.4	35.2	11.6	12.8	20.9	22.7	83.9	117.4	11.8
Std	17.9	38.0	53.0	69.5	34.2	15.1	29.6	33.7	23.3	14.6	13.9	20.4	16.4	66.1	80.6	16.5
	94.3%	94.3%	94.3%	86.7%	97.1%	94.3%	97.1%	98.1%	97.1%	95.2%	93.3%	98.1%	98.1%	66.7%	42.9%	98.1%

Table 6 Users age details

	Min	Max	Mean	Std
MCI	52	99	82.1	10.1
HC	45	95	63.8	14.7

Table 7 Classification results

	Precision	Recall	*F*-measure
HC	0.74	0.89	0.81
MCI	0.70	0.44	0.54
Avg	0.72	0.73	0.71

example of a mobile conversional agent successfully used in the context of AD has been recently reported in [48]. A purely automatic diagnostic tool paws the way of a quick instrument which enables mass screening of the population or even home training for improving cognitive abilities.

Another key feature of the proposed protocol is its extensibility, as the implemented tasks are decoupled from the logic of the application. Thanks to this, some other future developments can be addressed. First, additional standard tests may be digitized, e.g., the Frontal Assessment Battery (FAB) test [49], so that a more comprehensive framework for neurological assessment can be provided. Second, several other tasks could be investigated through the same equipment. For instance, house drawing has revealed to be discriminant for the AD/MCI discrimination [12]. Third, this work may be extended considering data acquired through different devices, for example, sensor-based smart gloves [50], and haptic technology [51], which have been successfully used for PD assessment. Fourth, handwriting-based measures recently showed promising results in investigating mental workload [17] as well as emotional states, such as depression and anxiety [52]. Hence, the battery of graphical tasks may be extended or refined for use in domains other than dementia assessment.

Funding This work has been funded by the Italian Ministry of Education, University and Research within the PRIN2015 - Handwriting Analysis against Neuromuscular Disease - HAND Project under Grant H96J16000820001.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

1. Cummings JL & Benson DF. Dementia: A clinical approach. Butterworth-Heinemann Medical; 1992.
2. Alzheimer's Association. 2017 Alzheimer's disease facts and figures. Alzheimers Dement. 2017;13(4):325–73.
3. Lang AE, Lozano AM. Parkinson's disease. N Engl J Med. 1998;339(15):1044–53.
4. Poewe W, Gauthier S, Aarsland D, Leverenz JB, Barone P, Weintraub D, et al. Diagnosis and management of Parkinson's disease dementia. Int J Clin Pract. 2008;62(10):1581–7.
5. Gauthier S, Reisberg B, Zaudig M, Petersen RC, Ritchie K, Broich K, et al. Mild cognitive impairment. Lancet. 2006;367(9518):1262–70.
6. Folstein MF, Robins LN, Helzer JE. The mini-mental state examination. Arch Gen Psychiatry. 1983;40(7):812–2.
7. Faundez-Zanuy M, Hussain A, Mekyska J, Sesa-Nogueras E, Monte-Moreno E, Esposito A, et al. Biometric applications related to human beings: there is life beyond security. Cogn Comput. 2013;5(1):136–51.
8. Lopez-de-Ipiña K, Alonso JB, Solé-Casals J, Barroso N, Henriquez P, Faundez-Zanuy M, … & Eguiraun H. On automatic diagnosis of Alzheimer's disease based on spontaneous speech analysis and emotional temperature. Cogn Comput. 2015;7(1):44–55.
9. Tseng MH, Cermak SA. The influence of ergonomic factors and perceptual-motor abilities on handwriting performance. Am J Occup Ther. 1993;47(10):919–26.
10. Impedovo D, Pirlo G. Dynamic handwriting analysis for the assessment of neurodegenerative diseases: a pattern recognition perspective. IEEE Rev Biomed Eng. 2019;12:209–20
11. Kawa J, Bednorz A, Stępień P, Derejczyk J, Bugdol M. Spatial and dynamical handwriting analysis in mild cognitive impairment. Comput Biol Med. 2017;82:21–8.
12. Müller S, Preische O, Heymann P, Elbing U, Laske C. Diagnostic value of a tablet-based drawing task for discrimination of patients in the early course of Alzheimer's disease from healthy individuals. J Alzheimers Dis. 2017;55(4):1463–9.
13. Müller S, Preische O, Heymann P, Elbing U, Laske C. Increased diagnostic accuracy of digital vs conventional clock drawing test for discrimination of patients in the early Course of Alzheimer's disease from cognitively healthy individuals. Front Aging Neurosci. 2017;9.
14. Schröter A, Mergl R, Bürger K, Hampel H, Möller HJ, Hegerl U. Kinematic analysis of handwriting movements in patients with Alzheimer's disease, mild cognitive impairment, depression and healthy subjects. Dement Geriatr Cogn Disord. 2003;15(3):132–42.
15. Slavin MJ, Phillips JG, Bradshaw JL, Hall KA, Presnell I. Consistency of handwriting movements in dementia of the Alzheimer's type: a comparison with Huntington's and Parkinson's diseases. J Int Neuropsychol Soc. 1999;5(1):20–5.
16. Bidet-Ildei C, Pollak P, Kandel S, Fraix V, Orliaguet JP. Handwriting in patients with Parkinson disease: effect of L-dopa and stimulation of the sub-thalamic nucleus on motor anticipation. Hum Mov Sci. 2011;30(4):783–91.
17. Drotár P, Mekyska J, Rektorová I, Masarová L, Smékal Z, Faundez-Zanuy M. Analysis of in-air movement in handwriting: a novel marker for Parkinson's disease. Comput Methods Prog Biomed. 2014;117(3):405–11.
18. Drotár P, Mekyska J, Rektorová I, Masarová L, Smékal Z, Faundez-Zanuy M. Decision support framework for Parkinson's disease based on novel handwriting markers. IEEE Trans Neural Syst Rehabil Eng. 2015;23(3):508–16.

19. Ünlü A, Brause R, Krakow K. Handwriting analysis for diagnosis and prognosis of Parkinson's disease. In: International Symposium on Biological and Medical Data Analysis. Berlin Heidelberg: Springer; 2006. p. 441–50.
20. Smits EJ, Tolonen AJ, Cluitmans L, van Gils M, Conway BA, Zietsma RC, et al. Standardized handwriting to assess bradykinesia, micrographia and tremor in Parkinson's disease. PLoS One. 2014;9(5):e97614.
21. Ponsen MM, Daffertshofer A, Wolters EC, Beek PJ, Berendse HW. Impairment of complex upper limb motor function in de novo Parkinson's disease. Parkinsonism Relat Disord. 2008;14(3):199–204.
22. Drotár P, Mekyska J, Rektorová I, Masarová L, Smékal Z, Faundez-Zanuy M. Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease. Artif Intell Med. 2016;67:39–46.
23. Alonso-Martinez C, Faundez-Zanuy M, Mekyska J. A comparative study of in-air trajectories at short and long distances in online handwriting. Cogn Comput. 2017;9(5):712–20.
24. Rosenblum S, Samuel M, Zlotnik S, Erikh I, Schlesinger I. Handwriting as an objective tool for Parkinson's disease diagnosis. J Neurol. 2013;260(9):2357–61.
25. Rosenblum S, Engel-Yeger B, Fogel Y. Age-related changes in executive control and their relationships with activity performance in handwriting. Human Movement Sci. 2013;32:363–76.
26. Luria G, Kahana A, Rosenblum S. Detection of deception via handwriting behaviors using a computerized tool: toward an evaluation of malingering. Cogn Comput. 2014;6(4):849–55.
27. Rosenblum S, Luria G. Applying a handwriting measurement model for capturing cognitive load implications through complex figure drawing. Cogn Comput. 2016;8(1):69–77.
28. Badarna M, Shimshoni I, Luria G, Rosenblum S. The importance of pen motion pattern groups for semiautomatic classification of handwriting into mental workload classes. Cogn Comput. 2018:1–13.
29. De Stefano C, Fontanella F, Impedovo D, Pirlo G, Scotto di Freca A. Handwriting analysis to support neurodegenerative diseases diagnosis: a review. Pattern Recogn Lett. 2019;121:37–45.
30. Impedovo D, Pirlo G, Angelillo MT, Vessio G. Towards a digitized protocol based on handwriting for the assessment of neurodegenerative disorders. CEUR Workshop Proc. 2017;1982:1–8.
31. Pirlo G, Cuccovillo V, Diaz-Cabrera M, Impedovo D, Mignone P. Multidomain verification of dynamic signatures using local stability analysis. IEEE Transactions on Human-Machine Systems. 2015;45(6):805–10.
32. Borson S, Scanlan J, Brush M, Vitaliano P, Dokmak A. The Mini-Cog: a cognitive 'vital signs' measure for dementia screening in multi-lingual elderly. Int J Geriatr Psychiatry. 2000;15(11):1021–7.
33. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist. 1969;9(3):179–86.
34. Reitan RM. Validity of the trail making test as an indicator of organic brain damage. Percept Mot Skills. 1958;8(3):271–6.
35. Della Sala S, Laiacona M, Spinnler H, Ubezio C. A cancellation test: its reliability in assessing attentional deficits in Alzheimer's disease. Psychol Med. 1992;22(4):885–901.
36. Dahmen J, Cook D, Fellows R, Schmitter-Edgecombe M. An analysis of a digital variant of the Trail Making Test using machine learning techniques. Technol Health Care. 2017;25(2):251–64.
37. Teulings HL, Contreras-Vidal JL, Stelmach GE, Adler CH. Parkinsonism reduces coordination of fingers, wrist, and arm in fine motor control. Exp Neurol. 1997;146(1):159–70.
38. Pirlo G, Diaz M, Ferrer MA, Impedovo D, Occhionero F, Zurlo U. Early diagnosis of neurodegenerative diseases by handwritten signature analysis. In: International Conference on Image Analysis and Processing. 2015. p. 290–297.
39. Pereira CR, Weber SA, Hook C, Rosa GH, Papa JP. Deep learning-aided Parkinson's disease diagnosis from handwritten dynamics. In: 29th SIBGRAPI conference on graphics, patterns and images. 2016. p. 340-346.
40. San Luciano M, Wang C, Ortega RA, Yu Q, Boschung S, Soto-Valencia J, et al. Digitized spiral drawing: a possible biomarker for early Parkinson's disease. PLoS One. 2016;11(10):e0162799.
41. Pullman SL. Spiral analysis: a new technique for measuring tremor with a digitizing tablet. Mov Disord. 1998;13(S3):85–9.
42. Impedovo D, Pirlo G, Mangini FM, Barbuzzi D, Rollo A, Balestrucci A, … & Plamondon R. Writing generation model for health care neuromuscular system investigation. In International Meeting on Computational Intelligence Methods for Bioinformatics and Biostatistics. Cham: Springer; 2013. p. 137–148
43. Werner P, Rosenblum S, Bar-On G, Heinik J, Korczyn A. Handwriting process variables discriminating mild Alzheimer's disease and mild cognitive impairment. J Gerontol B Psychol Sci Soc Sci. 2006;61(4):228–36.
44. Garre-Olmo J, Faúndez-Zanuy M, López de Ipiña K, Calvó-Perxas L, Turró-Garriga O. Kinematic and pressure features of handwriting and drawing: preliminary results between patients with mild cognitive impairment, Alzheimer disease and healthy controls. Curr Alzheimer Res. 2017;14(2017):960–8.
45. Brooke J. SUS-A quick and dirty usability scale. Usability Evaluation in Industry. 1996;189(194):4–7.
46. Boren T, Ramey J. Thinking aloud: reconciling theory and practice. IEEE Transactions on Professional Communication. 2000;43(3):261–78.
47. Giancardo L, Sanchez-Ferro A, Arroyo-Gallego T, Butterworth I, Mendoza CS, Montero P, … & Estépar RSJ. Computer keyboard interaction as an indicator of early Parkinson's disease. Sci Rep. 2016;6.
48. Griol D, Callejas Z. Mobile conversational agents for context-aware care applications. Cogn Comput. 2016;8(2):336–56.
49. Dubois B, Slachevsky A, Litvan I, Pillon BFAB. The FAB A frontal assessment battery at bedside. Neurology. 2000;55(11):1621–6.
50. Niazmand K, Tonn K, Kalaras A, Fietzek UM, Mehrkens JH & Lueth TC. Quantitative evaluation of Parkinson's disease using sensor based smart glove. In Computer-Based Medical Systems (CBMS), 2011 24th International Symposium. IEEE; 2011. p. 1-8
51. Li KY, Chu PY, Pickett KA. The effect of dopaminergic medication on joint kinematics during haptic movements in individuals with Parkinson's disease. Behav Neurol. 2017;2017:1–8.
52. Likforman-Sulem L, Esposito A, Faundez-Zanuy M, Clémençon S, Cordasco G. EMOTHAW: a novel database for emotional state recognition from handwriting and drawing. IEEE Transactions on Human-Machine Systems. 2017;47(2):273–84.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.