UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE INFORMATICA

Exploring the Usage of Text-Entry as a Digital Endpoint in Parkinson's Disease

Ana Morais

Mestrado em Informática

Dissertação orientada por:
Professor Doutor Tiago João Vieira Guerreiro
Doutor André Filipe Pereira Rodrigues

2022

Resumo

As doenças neurodegenerativas são caracterizadas pela destruição de neurónios vulneráveis de sistemas anatómicos específicos [19], onde as doenças identificadas como sendo mais comuns são as doenças de Parkinson, Alzheimer, Huntington e Esclerose Lateral Amiotrófica [45]. Embora com características variáveis entre doenças, têm sintomas semelhantes como declínio cognitivo e motor [12]. Com o avanço da tecnologia, esta tem sido explorada para auxiliar no diagnóstico e monitorização da condição neurológica de pacientes afetados por estas doenças.

A doença de Parkinson caracteriza-se pela perda de neurónios dopaminérgicos, e tende a manifestar-se através da rigidez e lentificação de movimentos, tremor, dificuldade em andar e postura irregular [3]. A doença de Alzheimer destaca-se pela perda de neurónios associados a funções cognitivas como memória, aprendizagem e pensamento, e tende a manifestar-se pela perda de memória e dificuldade na linguagem [5]. A doença de Huntington é causada por uma mutação num cromossoma e os sintomas comuns são a coreia e a bradicinesia [49]. Já a Esclerose Lateral Amiotrófica é caracterizada pela modificação de neurónios motores no cérebro e medula espinhal e consiste na paralisia e atrofia progressiva dos músculos [8]. Embora incuráveis, existe medicação para aliviar os sintomas, contudo o maior esforço tende a ser para garantir a melhor qualidade de vida possível aos pacientes.

A utilização de tecnologia para o diagnóstico e monitorização da evoluçã da doença tem vindo a crescer, particularmente com a utilização de sensores [30], embora a tendência seja no sentido de utilizar dispositivos diariamente utilizados, como por exemplo smartphones, para este efeito [18] [42]. Através de smartphones é possível fazer recolha de métricas de texto para analisar a performance de doentes, tanto por recolha ativa ou passiva [46] [47].

No sentido de sistematizar a investigação sobre doenças neurodegenerativas e a utilização de métricas de entrada de texto para o seu diagnóstico e monitorização, foi conduzida uma revisão sistemática. Dos 62 artigos identificados com potencial interesse, 48 eram provenientes de bases de dados e 14 de referências de referências. Após a remoção de duplicados e aplicação dos critérios de inclusão, nove foram incluídos na revisão sistemática. A análise demonstrou que a maioria dos estudos se foca na utilização de métricas de entrada de texto para a doença de Parkinson, embora haja também investigação noutras doenças neurodegenerativas como na Esclerose Múltipla. Em termos da duração do estudo e da colheita das métricas, pouca informação é fornecida, e quando mencionam, são díspares entre si. A revisão sistemática permitiu perceber também que o maior foco das investigações tem sido no diagnóstico precoce das doenças e não tanto na monitorização das mesmas. No que toca aos participantes dos estudos, a média de idades dos pacientes é de 63 anos, enquanto que dos controlos saudáveis é de 52 anos. Com uma incidência superior na doença de Parkinson, a maioria dos artigos utilizou como base de teste a escala UPDRS. Já no que toca a métricas recolhidas, a revisão demonstrou pouca variedade visto que as mais presentes são o hold time (tempo de espera), flight time (tempo de voo) e pressure (pressão).

Tendo em conta a não inclusão de clínicos no desenho de ferramentas de auxílio nos estudos analisados pela revisão sistemática e com o intuito de quebrar a tendência, foi conduzida uma sessão de trabalho em conjunto com o Centro Neurológico Sénior (CNS) para perceber do ponto de vista dos clínicos, o potencial da entrada de texto no diagnóstico e monitorização de doenças neurodegenerativas. Com uma abordagem inicial de apresentação de cinco métricas recolhidas pelo WildKey [46], sendo elas palavras por minuto, flight time, hold time, taxa de erros corrigidos e taxa de erros não corrigidos, foi pedido aos clínicos que estabelecessem conexões entre estas e conceitos de interesse. Foram identificados por eles doze conceitos, sendo eles bradicinesia, discinesia, apraxia, flutuações motoras e não-motoras, bradipsiquismo, defeitos cognitivos, dificuldade com tarefas duplas, ansiedade, tremor, cansaço, sonolência e movimentos involuntários. Depois desse primeiro exercício, foi pedido que relacionassem os conceitos encontrados com outras dezenove métricas recolhidas pelo WildKey [46] e a conclusão dos clínicos foi que seriam as mesmas conexões. Adicionalmente, foi possível extrair expectativas em relação aos dados, no que toca aos pacientes e aos seus estados ON e OFF, como por exemplo, que um paciente escreveria menos palavras por minuto do que um controlo saudável enquanto estaria sob o efeito da medicação (em ON), mas ainda assim mais do que enquanto estivesse fora do efeito da medicação (em OFF).

Já no que toca à visualização dos dados, houve uma preferência para os gráficos de linhas, onde o eixo dos xx corresponderia ao tempo em horas, e o eixo dos yy a cada uma das métricas a analisar. O conjunto de dados de semanas seriam então agregados com uma evolução por hora e idealmente com linhas verticais que indicassem momentos chaves do dia, como por exemplo, a altura onde a medicação fizesse mais efeito. Foi também abordada a diferença de necessidades expectável pelos clínicos de acordo com os pacientes a serem avaliados, e de acordo com os clínicos, seria expectável que uma pessoa com uma fase da doença mais avançada necessitasse de mais dados como auxílio, visto que com a progressão da doença ou até mesmo da idade, os pacientes tendem a não lembrar se se sentiram mal, se tomaram a medicação ou se falharam, e em que dias, por exemplo. Enquanto que um paciente com sintomas ligeiros e com uma vida ativa ainda consegue reportar esses detalhes. Idealmente, para uma avaliação ponderada, os clínicos necessitam pelo menos de uma semana de dados a terminar no dia da consulta de avaliação. Contudo, com a possibilidade de um maior conjunto de dados, os clínicos acreditam que seria interessante explorar a identificação da doença apenas pela utilização regular de dispositivos de introdução de texto.

Embora este estudo aborde principalmente a utilização de introdução de texto através de smartphones para a doença de Parkinson, visto ser a expertise dos clinicos envolvidos, e a sugestão de apresentação de dados ser introdutória, permite aos investigadores compreender de uma maneira sistematizada o que tem sido desenvolvido bem como as lacunas do trabalho relacionado, e que possam agora explorar e validar as conexões extraídas da sessão com os clínicos entre as métricas e os conceitos de interesse.

Palavras chave: doenças neurodegenerativas, doença de Parkinson, entrada de texto, diagnóstico, monitorização

Abstract

Neurodegenerative diseases are a group of diseases characterised by the loss of neurons and tend to be fatal. The most researched being Parkinson's disease, some connections have been established between this disease and the use of text-entry towards its diagnosis and monitoring. With such scattered information regarding neurodegenerative diseases and text-entry, a systematic review was carried out to show which diseases have been researched in that direction, being mainly PD but also MCI and MS. The main metrics collected were flight time, hold time and pressure. As previous research did not include clinicians participation towards the design of diagnosing and monitoring tools, this dissertation went a step further and worked together with clinicians to understand their expectations on data and its visualisations. Clinicians believe that text-entry does have potential towards the diagnosis and monitoring of neurodegenerative diseases. Clinicians also provided concepts of interest against recently suggested metrics, such as apraxia, bradykinesia and dyskinesia. Finally, it was possible to understand how clinicians would deem to be the best way to view the data for the patients' assessments.

Keywords: neurodegenerative disease, Parkinson's disease, text-entry, diagnosis, monitoring

For my family.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Parkinson's disease is a common progressive disorder which causes the loss of neurons [13]. PD is mainly defined by the difficulty in motor control and cognitive changes, which affect the patients' quality of life as it disturbs all panes of life [13]. Although the disease itself is incurable [42], there is medication which provides symptomatic relief, including Levodopa [13]. However, since it is a silent disease in its early stages, by the time the disease shows motor symptoms, part of the affected neurons are already lost [28]. The usage of smart devices in healthcare is an emerging area of research fuelled by the wide availability of such devices [38]. Since PD patients have impairments in coordination and rhythm stability, keystroke dynamics (such as keystroke timings), have become interesting sources of information, as a product of fine motor skills [28].

1.1 Motivation

Parkinson's disease is one that affects both mind and body, and can be very debilitating to its patients, as it tends to affect both cognition and motor abilities, having a tremendous impact in patients' life. As its presence can't be found through common diagnostic methods, i.e. blood samples, it can go undiscovered for quite a long time. Besides the diagnosis, it is also difficult to understand the patients' evolution since the main methods for monitoring are the patients' reports and assessments that depend on the clinicians' training and experience. However, there has been growing research towards founding new ways for diagnosis. Researchers have been trying to find alternatives, from using computers [22], games [51], sensors [30], to smartphones, to distinguish a Parkinson's disease patient from a healthy control. From these, wearable sensors have been the most researched, and whilst this shows great progress, there is still not a lot of research, particularly towards the monitoring of the disease through mainstream devices such as smartphones.

1.2 Context

This thesis was developed within LASIGE (Large-Scale Informatics Systems Laboratory), an investigation centre focused in Computer Science and Engineering where a bigger project focused on providing more and better quality data to everyone involved with the disease, both patients and clinicians, and with a partnership with CNS (Campus Neurológico Sénior). LASIGE has contributed with research as well as with tools, such as WildKey, which is a privacy-aware keyboard toolkit for data collection in-the-wild, provides the opportunity to collect patients' data in both implicit and explicit ways. Which means patients can contribute towards their health monitoring by performing specific tasks and/or simply by using their devices from the comfort of their own homes. WildKey collects several different metrics, from speed to errors and touch dynamics, among others. CNS is a rehabilitation centre specialised in Parkinson's disease among other neurological diseases that aims to improve the patients' life quality, by working with multidisciplinary teams.

1.3 Research goals

The main goal of the project is to find a way to help both patients and clinicians with the monitoring of neurodegenerative diseases, more specifically, Parkinson's disease, through the use of everyday technology - smartphones, and data collection - text-entry. For this, it is extremely important to understand what the clinicians' points of view are regarding the metrics they believe would be useful to them, and how they would like to see the data represented.

1.4 Approach

To better understand the disease, research was carried out to understand what has been done so far, and naturally, for one to understand where Parkinson's disease fits, it makes sense to look at neurodegenerative diseases in general. As the search revealed a few articles and no systematic reviews, it seemed natural to close this gap by developing a systematic review on the topic. After understanding the current state of the art, the focus would be redirected towards the clinicians and their needs. So, a study to understand the potential of text entry with clinicians was conducted with members from the CNS. From the literature review and study with clinicians, I derived a set of directions for future research.

1.5 Contributions

As the main contributions, this dissertation aims to provide knowledge for future research by offering an identification of the current research on neurodegenerative diseases on diagnosing and monitoring through text-entry on smartphones, through a systematic review, and a characterisation of clinicians' needs for reporting, so that other researchers can look for clinical utility in using text-entry as a disease endpoint. Additionally, it offers a potential clinical connection between text-entry metrics and concepts of interest for the disease, which can now be explored and validated. Finally, it offers a set of recommendations for dashboard design that allows the use of this data in a clinical environment.

1.6 Document's Structure

The document is organised as follows.

Chapter 2 - Background

In this chapter a background on neurodegenerative diseases, sensors and text-entry is provided. Four neurodegenerative diseases were listed and explained according to research, being the most common diseases: Parkinson's disease, Alzheimer's disease, Huntington's disease, and Amyotrophic Lateral Sclerosis.

Chapter 3 - Touch Typing as a Digital Endpoint for Neurodegenerative Diseases: a Systematic Review

This section shows the work carried out to review existing research on diagnosing and monitoring of neurodegenerative diseases through text-entry.

Chapter 4 - Understanding the potential of text-entry with clinicians

Through this chapter, you will understand the clinicians points of view regarding diagnosing and monitoring of neurodegenerative diseases through text-entry, on their expectations regarding the data as well as how they deem to be the ideal way to represent it.

Chapter 5 - Discussion

In this section, an overview is given of the results collected by this dissertation, through both the systematic review and work study with clinicians.

Chapter 6 - Conclusions

In this final chapter, the final conclusions are presented, the main benefits and limitations are discussed and future work is suggested.

Chapter 2

Background

In an initial search on Parkinson's disease, there were quite a few articles on proposed methodologies for the identification of the disease, through various technologies. After noticing the amount of articles, it was concluded that it would be beneficial to understand how the diagnosis and monitoring was being done for neurodegenerative diseases. This new search helped understand the bigger picture on neurodegenerative diseases as well as how Parkinson's disease fits within them, and how technology could help.

2.1 Neurodegenerative diseases

Neurodegenerative diseases represent a large section of neurological disorders [45]. These are characterised by the loss of vulnerable neurons [19] in specific functional anatomic systems. The reason for their appearance is yet unknown and they tend to progress relentlessly [45]. With hundreds of different diseases in this category, the most prominent diseases are Parkinson's disease (PD), Alzheimer's disease (AD), Huntington's disease (HD) and Amyotrophic Lateral Sclerosis (ALS) [45]. Even though the characteristics vary according to each disease, there are shared symptoms like a progressive cognitive decline, motor dysfunctions, and deficits in gait and balance [12]. Quantitative measurements of performance can provide clinicians with critical information regarding the patient's disease severity and progression, and help in the long term with improving the patient's quality of life. In recent years, there have been advances towards the use of technology for providing measures of mobility performance to help understand the patient's neurological condition [12].

2.1.1 Parkinson's Disease

Parkinson's disease is characterised as a chronic neurological disorder which causes progressive disability due to the loss of dopaminergic neurons [13]. Being the second most common neurodegenerative disorder, presents an annual incidence rate of 8-18 per 100,000 persons [13]. PD is mainly defined by motor impairment, involving rigidity, gait impairment, postural instability, tremor, and bradykinesia. However, non-motor symptoms also characterise this disease, including cognitive and mood changes or sleep disturbances. This symptom diversity affects the patients' quality of life as it disturbs the physical, mental, and social panes of life [3]. Subtle motor manifestations can appear years before the clinical diagnosis, and they often go unnoticed particularly in the early stages of the disease. After the diagnosis, patients usually progress towards severe disability and a shortened life span [16].

Although the disease itself is incurable [38], there is medication which provides symptomatic relief, including Levodopa [13], that when administered at an early stage improves the clinical outcomes [42]. However, there is some difficulty in adapting treatment as there is a lack of clear and objective methods that quantify and monitor the different disease stages [3]. Another challenge in treatment comes from the fact that by the time the motor signs are present, over 50% of neurons in the affected area have already been lost [16]. Besides medication, there are surgical interventions available through pallidotomy (the destruction of a tiny area of the brain) and Deep Brain Stimulation (DBS). Usually, DBS is a choice when the patient does not respond to Levodopa. Although DBS provides better results than Levodopa, therapy wise, patients that undergo DBS are more susceptible to serious adverse effects, such as fatal cerebral haemorrhage [23].

When it comes to the clinical evaluation of PD, the most used instrument is the Unified Parkinson's Disease Rating Scale (UPDRS). It is a standardised test which provides an overall score of the patients' functional capabilities [3]). This scale comprises four parts: I – non-motor experiences; II – Motor experiences; III – Motor examination; IV – Motor complications. Each part has several items, each item is scored from 0 (normal) to 4 (severe) and for each part, the sum of these scores makes the total score [35]. The UPDRS-III attempts to quantify the severity of PD's motor symptomatology [42], namely, rigidity, speech, facial expressions, resting tremors, among others [22]. Despite being the most accepted standard, it requires a well-trained and experienced clinician to make an evaluation with a low risk of misdiagnosis [3] [25] ; as well as attendance of the patient in the clinic, limiting the ease and frequency of administration. It has been shown that patients visit their neurologists every 2 to 6 months [11], this added to short clinical consultations (often less than 15 minutes) shows that clinicians are only able to take a 'snapshot' of the patient's condition rather than a continuous assessment [43]. There are also complementary methods to detect and quantify psychomotor dysfunction, such as the finger-tapping tests. The Alternating finger-tapping test (ATF) is one of the varieties, where the subject while using a single hand, alternatively presses two buttons as fast as possible for a period of time. The test is done for both hands and the final score is the average number of pressed keys, between the two [3]. As the early motor signs of PD are so mild, patients tend to ignore them, causing the condition to remain under the radar for a critical period of time. Hence the need for accessible tools which monitor subjects remotely, unobtrusively and throughout their daily routine [42].

The usage of smart devices in healthcare is an emerging area of research fuelled by the wide availability of such devices [42]. The mPower study is one of the largest studies where a smartphone activity tracker was used to collect longitudinal data from PD patients and controls, through touchscreen typing, memory, voice, posture and gait tests. The main limitation of this study was the reduced compliance as it required the subjects' active participation [8]. Other studies have captured data unobtrusively in the background during routine typing in smartphones, reflecting the user's natural behaviour. Since PD patients have impairments in coordination and rhythm stability, keystroke dynamics (such as keystroke timings), become interesting sources of information, as a product of fine motor skills [28].

2.1.2 Alzheimer's

Alzheimer's disease is a type of degenerative brain disease with an increased prevalence of 89% since 2000 [31]. It is believed that by the time the symptoms arise, AD has already begun 20 years ago, as the person does not notice the small changes until after the symptoms become noticeable and manifest themselves, being either memory loss or language problems [5]. AD is characterised by the progressive destruction of neurons involved with cognitive functions like memory, thinking and learning. However, as the disease advances further, neurons in other parts of the brain also become damaged or even destroyed, having a very significant impact on basic bodily functions like walking and swallowing, affecting the patient's ability to carry out everyday activities [5]. Being ultimately fatal, it is considered to be the sixth leading cause of death in the United States [31]. Existing research identifies three stages in Alzheimer's, the first one being preclinical AD, the second mild cognitive impairment (MCI) due to AD and lastly, dementia due to AD, where in the later two, symptoms are present but with varying degrees of impact [5].

When it comes to treatment, even though there are six drugs approved, as they temporarily improve the symptoms through an increase in neurotransmitters in the brain [5], it is generally accepted that the likelihood of reversal of anatomic and physiological changes, such as neuronal death, massively decreases with disease progression. With this scenario, attention rises towards the early diagnosis of AD for future studies [31]. Additionally, there are some non-pharmacological therapies that aim to improve AD patients' lives by maintaining or improving their cognitive functions, as well as reducing behavioural symptoms like agitation, depression, agression and apathy [5].

Without routine diagnostic tools for early detection of AD [15], physicians use a variety of approaches and tools, together with multidisciplinary specialists to help with the diagnosis, typically starting with the family history, conducting cognitive tests, physical and neurological examinations and blood tests to rule out potential causes of dementia symptoms like tumours and vitamin deficiencies [5]. However, some believe that it is imperative to determine the root causes of AD as well as find a clinical presentation to aid in the early diagnosis [15], and that the natural research progression will be towards using digital biomarkers, by leveraging the widely available wearable technologies and mobiles as they can provide immediate access to information with an extremely low impact on the healthcare system [31].

2.1.3 Huntington's Disease

Huntington's disease is a rare hereditary neurodegenerative disorder caused by a mutated chromosome [49] and is associated with psychiatric and cognitive symptoms [36]. Although it is most common amongst the fourth or fifth decade of life, the first symptoms can manifest themselves anywhere from childhood to older age, typically being lethal within fifteen to twenty years [49]. In early stages, HD tends to manifest itself through behaviour disturbances and learning difficulties at school. The most common sign of HD is chorea [48], characterised by random dance-like movements [53], which typically spreads gradually to all muscles and with its progression, there is also a retardation of the psychomotor processes [48]. Patients tend to experience additional symptoms such as bradykinesia, which is the slowness of movement, akinesia, which is the difficulty in starting movements, dystonia, which is characterised by rigid slow movements which lead to an abnormal posture [48], and cognitive impairment, affecting attention and emotion recognition [49], having a great impact on the patients' daily routines. As for treatment, research shows that there is still no treatment that can prevent or slow down HD, so the focus ends up being on improving the patients' quality of life as much as possible through medical and non-medical means [6]. However, due to its nature, there is hope and it is believed that HD may be, perhaps, one of the most treatable neurodegenerative diseases [36].

HD is usually diagnosed through family history, if available, and in most cases, through genetic testing [7]. Even though there are additional tests which can support the diagnosis, mainly by ruling out other possibilities, usually they are not necessary. As for the motor impairment onset, clinicians use as baseline test the UHDRS (Unified Huntington's Disease Rating Scale), with a 99% confidence that signs are attributable to HD. A total motor score of approximately 15 in an adult who experiences symptoms like delayed eye movement, chorea and difficulty in walking, tends to be a very supportive indicator of the diagnosis [49]. The future of HD seems to be mainly focused on the search of biomarkers and treatment to stop the onset and effectively stop the disease altogether [48].

2.1.4 Amyotrophic Lateral Sclerosis

Amyotrophic Lateral Sclerosis is a neurodegenerative disease characterised by its progressive paralysis, as it affects both the motor neurons in the brain and the spinal cord. With an increased incidence and prevalence as people age, it tends to start with an insidious focal weakness and then progress to involve most muscles, so badly that typically the patient's death is caused by respiratory paralysis within three to five years [10]. Symptoms wise, patients tend to experience muscle stiffness and spasticity, where the stiffness of the muscle may be associated with discomfort or pain, as well as spontaneous muscle twitching and atrophy, which is the loss of muscle tissue. Besides the motor impact, ALS patients also tend to develop dementia due to the neuronal degeneration [10].

Research shows that there is no therapy available that provides a substantial clinical benefit for ALS patients, however, there are drugs which help by suppressing excessive motor neuron firing and stress [10]. Additionally, there has been growing research towards the use of stem cells to help protect surviving motor neurons [40]. As for diagnosis, there is not a definitive test for ALS so it tends to be based on clinical investigation to exclude other causal possibilities for the presenting symptoms and it also requires evidence of disease progression [24].

2.2 Monitoring with sensors

With the advancement of wearable technology, it has been made possible to monitor individuals' health outside of the clinic. Sensors have been designed to help understand patients' disease progression as well as with clinical care [9] [30] [34]. Remote monitoring can be extremely useful as it allows the detection of sporadic behaviours that do not show during clinical appointments or even behaviours that present themselves differently in the clinic and at home [30]. Different types of sensors have been found in literature for real-time data gathering, being classified in two categories, wearable and ambient sensors. Whilst the former are usually placed on patient's bodies, with the data being captured and then transmitted through wireless connections, the latter are placed in the environment hence not being required to be used directly by patients. There is, however, a third category which is created by a combination of these together, creating hybrid systems [12].

Smartphones, being one of the technologies accessible by most, have evolved to have built-in sensors, the most well known being accelerometers, gyroscopes and GPS. However, research has shown that smartphones can also help with assessing depression, influenza and heart diseases [1]. The great advantage of mobile technology for assessing and monitoring is that it can be easily integrated into the patient's life since it is simple, user-friendly, comfortable and non-invasive [18]. Additionally, despite the technology advancements, it still is up to the patient to accept and adhere to wearing sensors during daily activities [30]. However, smartphones can help mitigate this as it is an already present device in most people's lives [18]. With relation to neurodegenerative diseases, research has found that there is potential in sensors to help with assessing the patients' functional state, disease progression through symptoms like tremor [44], as well as response to therapy [4].

2.3 Text-entry

With the increased use of smartphones, more data is being generated and text input is one of the most common tasks, as it is involved in texting, emailing and social networking [47]. This sparked a growing interest in leveraging the everyday typing behaviour data for research purposes, across different fields, such as keyboard performance improvement, to increase speed and accuracy, biometrics, for authentication methods, linguistics, for communication through computers, and health, for early disease detection [46]. Two ways to study typing behaviours were found, one through experience sampling and one through passive sensing. Experience sampling prompts users to interact by performing specific tasks, whilst passive sensing collects data from the background whilst the patient uses the device during daily life [46]. Wildkey is a toolkit created for text entry studies that allows both data collection types, and that can provide text entry metrics on speed, error rates and touch behaviours [47].

Chapter 3

Touch Typing as a Digital Endpoint for Neurodegenerative Diseases: a Systematic Review

With the understanding of neurodegenerative diseases as well as the potential technology, an interest developed towards identifying what research had already been done for the use of text entry towards the identification and monitoring of neurodegenerative diseases. Given the amount of research available and the inexistence of a systematic review on the topic, the opportunity of creating one, revealed itself.

3.1 Introduction

Even though there is no cure for neurodegenerative diseases, research is being carried out towards this goal, mainly through neuronal regeneration [17] [37]. For the meantime there are treatments aiming for symptom relief [14]. When it comes to diagnosis, providing adequate training to clinicians is imperative for an accurate diagnosis [20]. Additionally, the early signs of the conditions are often ignored by patients, leaving the disease off the radar during critical time periods. This leads to the need of an accessible tool, able to monitor subjects remotely and send an alert to see a clinician if symptoms arise [42].

The use of personal electronics by the general population has brought to interest the users' data from natural interaction with the devices [2]. The data collected can be used for many purposes, including health monitoring, with the potential for early disease detection [46]. One of the most significant examples of this is the mPower study with over 9,000 participants, where a smartphone activity tracker was used to collect longitudinal data from PD patients and controls, through touchscreen typing, memory, voice, posture and gait tests. Despite its success, the study required the active participation of its users which had the unintended consequence of high dropt-out rates [28]. There is, however, another possible approach that might mitigate this issue, which is through passive sensing [46].

In order to identify and summarise the published literature on text entry towards the assessment/diagnosis and monitoring of neurodegenerative diseases, particularly with virtual keyboards, a systematic review was carried out: Touch Typing as a Digital Endpoint for Neurodegenerative Diseases: a Systematic Review.

3.2 Methods

The systematic review of the literature was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA 2020 guidelines) [41]. A protocol for the review was registered and can be accessed at PROSPERO [CRD42022293379].

3.2.1 Search strategy

Electronic database search was performed in April 2022, on the MEDLINE/PUBMED, the Institute of Electrical and Electronics Engineers (IEEE), and the Association for Computing Machinery (ACM) databases. The search was conducted with the following keyword combinations: (Huntington OR HD OR Parkinson OR PD OR Alzheimer OR AD OR Amyotrophic Lateral Sclerosis OR ALS OR Neurodegenerative OR Tremor OR Dementia OR Cognition) AND (text-entry OR text-input OR text entry OR text input OR typing OR touch input) AND (smartphone OR touchscreen OR mobile OR tablet OR Virtual Keyboard) for both title and abstract.

3.2.2 Eligibility Criteria

Articles were evaluated and selected according to the following eligibility criteria: Study must comprise (1) neurodegenerative patients (i.e., population type), (2) use finger-typed text entry (i.e., methodology), (3) published in a peer-Review journal or conference, (4) between 2007 and 2022, (5) in English language. The articles were assessed independently by two evaluators. Disagreements were solved by a third senior member of the research team.

3.3 Results

3.3.1 Studies included

Of the studies researched, 62 were identified as of potential interest. 48 from database searching and 14 from references of references. Six duplicate articles were removed, leaving 56 to be screened based on title and abstract eligibility. Out of these, 31 studies did not comprise neurodegenerative patients and 16 had no relation to finger-typed text entry on touch keyboards, and were therefore excluded for not meeting the inclusion criteria. Nine articles were included in the systematic review (please, see Figure 1 - PRISMA flow diagram).

3.3.2 Studies Duration

One article mentioned it was during a regular visit to the clinic [26], two mentioned a one day visit [3] [27], one specifies six months [39], one reported about five clinical visits with a three month interval, and the remaining four did not provide any information [25] [28] [42] [54]. This makes it possible to assume there is no standard duration for studies and the data collection time varies according to the number and duration of sessions.

3.3.3 Studies' Aims

Of nine articles, seven focus on Parkinson's disease [3] [25] [26] [27] [28] [42] [54], one in Mild Cognitive Impairment [39] and one in Multiple Sclerosis [32].

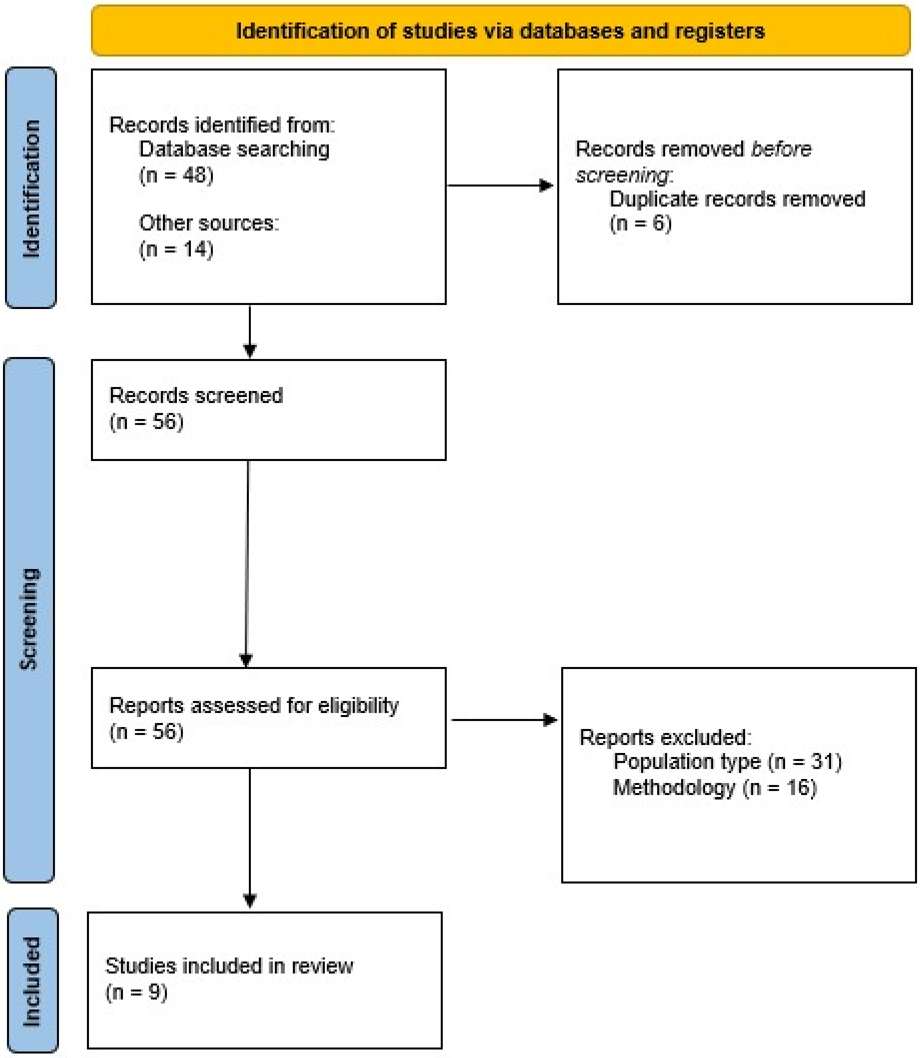

Figure 3.1: PRISMA flow diagram. PubMed/Medline = 19, IEEE = 15, ACM = 14. Reference of reference = 14.

Four articles focused on the creation of unobtrusive and accessible tools to monitor PD symptoms helping with the early diagnosis and with decreasing the gap between the disease onset and start of treatment [25] [26] [28] [42]. Two questioned whether touchscreen typing can be used as an indicator of disease, against baseline tests for PD and MS [27] [32]. One focused on an automated method to identify MCI individuals [39]. One develops a test for routine analysis of touchscreen typing for PD patients [3]. And another creates a prediction algorithm to tackle typing errors made by PD patients to improve their experience [54]. This shows that the majority of the research being carried out on the diagnosing and monitoring of neurodegenerative diseases, has been mostly focused on Parkinson's disease and its diagnosis.

3.3.4 Participants' demographics

From the nine studies included, 66% of studies used one dataset, 22% used two datasets and 11% used three datasets. From the nine studies analysed in the systematic review, a total of 346 participants were included. 215 participants were patients diagnosed with neurodegenerative disease (mean age 63 +/- SD, 50% men), and 131 participants were healthy controls (mean age 52 +/- SD, 44% men). The

Table 3.1: Data synthesis of the reviewed articles.

Author Year	Aim	Disease being assessed	Duration	Outcomes measured	Population size	Inclusion / exclusion criteria	In-clinic / free living	Metrics collected	Compared to other baseline tests
Iakovakis et al. 2019 [8]	Creating a passive and accessible tool; quantification of fine-motor skills' decline in early-PD; development of digital tools for remote symptom screening.	PD	duration of a regular visit at the clinic	CNN capacity to classify early-PD patients vs HC using raw FT and HT sequences	Total in--clinic: 33 Total PD patients: 18 Avg age: 61 yo Avg disease duration: 2.5 yrs Total HC: 15 Avg age: 57 yo Total free-living: 111	minimum 10 typing sessions; minimum 40 keystrokes/typing session	both	hold time and flight time	N/A
Papadopoulos et al. 2020 [14]	Development of an accessible PD screening tool with passive monitoring of subject-smartphone interaction searching for signs of PD; goal:reduce the gap between disease onset and treatment.	PD	N/A	Early detection of PD via multiple sources with multi-symptom analysis	Total in-clinic: 22 Total PD patients: 14 Avg age: 60.7 yo Avg disease duration: 7.5 yrs Total HC: 8 Avg age: 50.5 yrs Total free-living: 157 self-reported Total PD patients: 26 Avg age: 62.5 yo Total HC: 131 Avg age:54.5 yo	N/A	both	hold time, flight time, acceleration	MDS-UPDRS part III
Iakovakis et al. 2020 [7]	Remote and unobtrusive screening of subtle FMI via free-living smartphone usage; useful for consolidating early and accurate PD diagnosis.	PD	N/A	Cantouchscreen typing combined with deep learning provide an efficient tool for PD remote screening	Total free-living: 39 1st validation Total PD patients: 22 Avg age: 58.6 yo Avg disease duration: 2.5 yrs Total HC: 17 Avg age: 54.6 yo Total free-living: 26 2nd validation Total PD patients: 9 Total HC: 17	Minimum t 40 keystrokes/typing session	free-living	hold time and flight time	MDS-UPDRS part III

Arroyo-Gallego et al. 2017 [2]	Development of a home-based, high-compliance, and high-frequency PD motor test by analysis of routine typing on touchscreens	PD	One day visit	Correlation between the proposed multi-symptom algorithm and current motor tests	Total in-clinic: 51 Total PD patients: 24 - 3 (insufficient data) 21 Avg age: 59.2 yrs Total HC: 27 - 4 (insufficient data)) 23 Avg age: 54.35	20 keys per minute for at least half of the typing time	in-clinic	flight time	AFT
Iakovakis et al. 2018 [10]	Development of an accessible, non-intrusive tool for early-PD screening in a daily life setting; detection of fine motor skills' decline in early-PD via analysis of patterns emerging from finger interaction with touchscreen smartphones during free-living typing	PD	N/A	Identification of early-PD stage with a classification pipeline	Total in-clinic: 33 Total PD patients: 18 Avg age: 61 yo Avg disease duration: 2.5 yrs Total HC: 15 Avg age: 57	10 short text excerpts	in-clinic	hold time, flight time, normalised pressure	MDS-UPDRS part III
Ntracha et al. 2020 [13]	Creation of an automated method to identify patients with MCI using different sets of features and classifiers	MCI	6 months	Can a digital biomarker classifier distinguish MCI patients from HC	Total free-living: 23 Total MCI patients: 11 Avg age: 67.2 yo Established diagnosis within the last 3 years. Total HC: 12 Avg age: 66.2 yo.	Minimum t 40 keystrokes/typing session	free-living	hold time and flight time	MMSE, FRSSD, FUCAS, BDI, GDS.
Iakovakis et al. 2018 [9]	Exploitation of the best performing features from mobile touchscreen typing in a binary setting (HC vs PD) as predictors of specific MDS-UPDRS part III items, in order to ascertain symptom severity	PD	One day visit	Prediction of the MDS-UPDRS part III item scores in-clinic vs free-living	Total in-clinic: 33 Total PD patients: 18 Avg age: 61 yo Avg disease duration: 2.5 yrs Total HC: 15 Avg age: 57 yo Total free-living 210 (48 due to age range)	48 - 80 yo, more than 10 valid feature vectors to contain enough typing sessions	both	hold time, flight time, normalised pressure	MDS-UPDRS part III

Lam et al. 2021 [11]	Determine whether keystroke dynamics can be utilised as a biomarker in MS regarding disease activity, clinical disability and fatigue.	MS	5 clinical visits with a 3-month interval	Feasibility and reliability of free-living keystroke dynamics collected by smartphone technology to assess fatigue, MRI disease activity and clinical disability in MS	Total free-living: 103 Total MS patients: 85 Avg age: 46.4 yo Avg disease duration: 5.7 Total HC: 18 Avg age: 45.2 yo	conventional use/typing on a smartphone, 20 - 65 yo, no other relevant diagnoses or impairments	free-living	Press-Press Latency, Release-Release Latency, Hold Time, Flight Time, Pre-Correction Slowing, Correction Duration, Post-Correction Slowing, After Punctuation Pause	MRI, EDSS, (NHPT), SDMT, FSS and CIS-F.
Wang et al. 2021 [17]	Creating a prediction algorithm to improve PD users' experience by tackling errors	PD	N/A	Improve text entry experience for PD users through text suggestion	For the survey (online): Total PD patients: 16 Avg disease duration: 8.4 Typing behaviour study: Total in-clinic: 16 Total PD patients: 8 Avg age: 60.5 yo Total HC: 8 Avg age: 23.6 yo Prediction algorithm: Total PD patients: 8 Avg age: 59.9 yo	N/A	in-clinic	typing speed, character error rate, word-level error rate, keystroke per character, top-K accuracy, time interval of adjacent touch points, user experience and feedback	N/A

distribution within the cohorts would suggest that there is no specific tendency towards one gender being more prone to neurodegenerative diseases. Additionally, the age difference between patients and healthy controls averages by ten years.

Six articles mention the cohort's education levels. Around 68 (85%) of healthy controls have higher education compared to 12 (15%) who finished highschool. As for patients, 49 (30%) finished high school and 111 (69%) have higher education.

When it comes to baseline test ratings, only studies focused on Parkinson disease presented data, namely the UPDRS test [25] [27] [28] [42]. The average score was 18 points for PD patients and 1 point for healthy controls. As for years with the disease, four studies mentioned an average of 2.5 years since onset [25] [26] [27] [28], three mention the average of years since diagnosis being 5.7, 7.5 and 8.4 ([32] [42] [54] respectively), one study mentioned that the official diagnosis was done within the three previous years of the article [39], and one did not provide any information with this regard [3]. This makes it possible to assume there is no standard way of reporting length of disease. Another criteria evaluated in some articles was the smartphone usage of the users, only four articles mention this. Of these, three articles [26] [27] [28] mentioned that participants (e.g., patients) have an average of 3 years of smartphone usage, and one article [25] only refers that the majority of the participants have been using smartphones for more than a year.

3.3.5 Participants' inclusion and exclusion criteria and experimental setting

Only two articles specified an age related criteria for data collection (20 to 65 years [32], and 40 to 80 years [28]), three require an exact number of keystrokes (at least 40) [25] [26] [39], one requests a number of excerpts executed (10) [28], one mentions a rate of keys per minute (20, for at least half of the typing time) [3], and two do not disclose their inclusion or exclusion criteria [42] [54].

Out of the nine articles, three collected data in-clinic only [3] [28] [54], three collected data in free living only [25] [32] [39], three collected data in both scenarios [26] [27] [42].

3.3.6 Outcomes

As for outcomes, four of the nine articles were focused on the early diagnosis of PD patients, one through a Convolutional Neural Network (CNN) using raw flight time (FT) and hold time (HT) sequences [26], one study analysed the possibility of using a multi-symptom approach to detect PD [42], one study used deep learning for screening parkinsonian subtle fine-motor impairment [25] and one study used a classification pipeline for detecting fine motor skills decline in early PD patients [28]. In the remaining five articles different approaches were identified. For instance, one study concerned with identifying a correlation between the multi-symptom algorithm proposed and the current motor tests [3], one study aimed to test whether a digital biomarker classifier could distinguish MCI patients from HC [39], one study aimed to understand whether it was possible to predict the UPDRS III single item scores, and whether this worked both in clinic and in the wild [27], one study [32] aimed to understand whether keystroke dynamics could be used to assess fatigue, MRI disease activity and clinical disability in MS, and the last one [54] focused on improving the users' experience through text suggestion. From the outcomes collected, it is possible to understand that the research carried out so far has been mostly focused on the collection and validation of metrics, and no research has included clinicians to provide their standpoint.

3.3.7 Metrics

The metrics most commonly collected are the flight time (89%), hold time (78%) and pressure (22%). Other metrics include acceleration, press-press and release-release latencies, correction duration, after punctuation pause, character error rate, keystroke per character, and typing speed.

3.3.8 Baseline tests

57% of the articles focused on Parkinson's disease use UPDRS III as a baseline test, 11% use the AFT test, whilst the other make no mention of any baseline tests. One article related to Mild Cognitive Impairment mentions five different baseline tests: the Mini Mental State Examination (MMSE), the Functional Rating Scale for Symptoms of Dementia (FRSSD), the Functional Cognitive Assessment Scale (FUCAS), the Beck Depression Inventory (BDI) and Geriatric Depression Scale (GDS). As for the article on Multiple Sclerosis, six baseline tests are considered: the MRI, the EDSS, the NineHole Peg Test (NHPT), the Symbol Digit Modalities Test (SDMT), the Fatigue Severity Scale (FSS) and the Checklist Individual Strength Fatigue subscale (CIS-F).

3.4 Summary

This systematic review aimed to understand how text entry is being used towards assessing and monitoring neurodegenerative diseases. It helped finding that the majority of research being carried out is towards Parkinson's disease. The most common comparator for results validation is the UPDRS III test, that identifies Parkinson's disease. There seems yet not to exist a preference between in-the-clinic and in-the-wild data collection, although it is possible to see that in-the-wild collection is able to assess more people when compared to in-the-clinic collections, reaching bigger cohorts. The main metrics being collected are the hold time, flight time and pressure, showing a good correlation with symptomatic variation. As for the outcomes set, it has been shown that it is possible to distinguish disease patients from healthy controls with promising results, and when compared against baseline tests, one was found to outperform the existing other. As for prediction of test scores, one study found that there was a positive correlation between the performance of the dominant hand (for bradykinesia, hypokinesia and rigidity) and the UPDRS III item scores, holding true for both in-the-clinic and in-the-wild evaluations. Finally, it has also been shown that it is possible to improve patients' experience during their smartphone usage through text suggestion.

The assessed articles identified some limitations themselves. Between all articles, there were issues with the cohort size being too small, the suggested approach leading to false positives, having reduced medical transparency and loss of information due to data aggregation. Other concerns included the time in which patients had to refrain from medication being only 8 hours, when it usually takes 12 hours to reach the "practically off" state; comparing young healthy controls against older Parkinson patients instead of non-Parkinson's elders; the use of a PD dataset to train an algorithm to identify MCI patients; and potential privacy and security issues related to the written speech content.

With these findings in place, it was time to get the clinicians perspectives on text-entry metrics and how these can help with diagnosing and monitoring of neurodegenerative diseases.

Chapter 4

Understanding the potential of text-entry with clinicians

Previous studies regarding monitoring of neurodegenerative diseases have mostly included onsite appointments, which require both the patient's self-report as well as additional information, being provided either by a relative or a baseline test [33]. This may present a problem, since onsite evaluations only capture a snapshot of that particular moment. Considering the reality of many outpatient clinics where patients can be evaluated seldomly throughout the year, the clinical impression of disease progression may be harder to attain [33] which may impact therapeutic decisions, patient well-being and quality of life. Furthermore, patients' self-report can be biassed due to several reasons, including cognitive impairment and depressive symptoms or sleep disorders, making it difficult to ascertain fundamental details [33] [35].

There has been growing research on assessing symptoms and disease progression in a number of neurodegenerative diseases [47]. Monitoring through typing analysis, attempting to mitigate some of the issues referred above, may allow healthcare teams to possess important data regarding symptom worsening and/or disease progression, collected as a free-living evaluation during the patient's usual activities of daily living.

After the overview of the neurodegenerative diseases and the use of text-entry towards the diagnosis and monitoring of diseases done, and with the aim to understand the clinicians perspective on the matter and to provide knowledge for a reporting tool, a work session followed.

4.1 Research goals

The main goals for this study with professional stakeholders were to understand from the clinicians' point of view, how text-entry could help them with making better assessments of the patient's health progression, what their expectations would be for the data, and how it would make best sense to see the data represented.

4.2 Participants

Four participants were recruited: two clinicians, one speech therapist and one occupational therapist, all working closely with Parkinson's Disease Patients.

4.3 Procedure

The session took place at Campus Neurológico Senior (CNS) on the 5th of September 2022. During the one hour session, the participants were briefed about the research and aims of the study, followed by the explanation of the outline of the session. Five metrics from the Wildkey toolkit [47] were selected and presented, with descriptions, to get the participants to start the discussion without becoming overwhelmed. The five selected were:

- Words per minute - number of words written by the patient, per minute
- Flight time - time it takes for the patient to go from one key to another
- Hold time - time it takes for the patient to press a specific key
- Corrected error rate - percentage of corrected errors
- Uncorrected error rate - percentage of uncorrected errors

Once the metrics were presented, the participants were asked to identify concepts of interest and correlate them with the metrics. In a white board, the metrics were written and the participants placed notes with the relevant concepts around each of them.

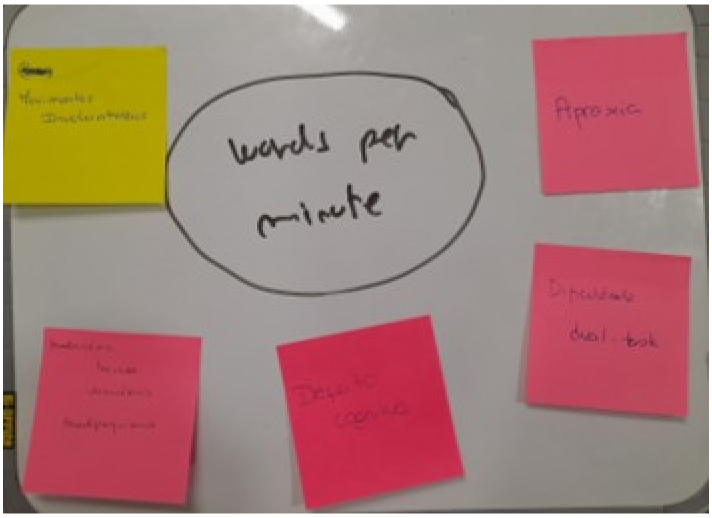

Figure 4.1: Example picture of the first exercise being carried out. Correlation between metrics and concepts of interest.

Following this exercise, the remaining nineteen metrics were presented with brief descriptions, by categories, to understand both the connection with the previously identified concepts of interest and, which could have potential interest to be analysed during a patient's review:

- Speed
  - Time per word

- Errors
  - Total error rate
  - Insertion error rate
  - Omission error rate
  - Substitution error rate
  - Error correction attempts
- Touch dynamics
  - Touch major / minor
  - Touch offset
  - Key selected
  - Motion info
- Action and character counts
  - Action count
  - Correction action count
  - Early action count
  - Number of auto corrects
  - Number of changed characters
  - Number of selected suggestions
  - Number of written character
  - Number of written numbers
  - Number of written special characters

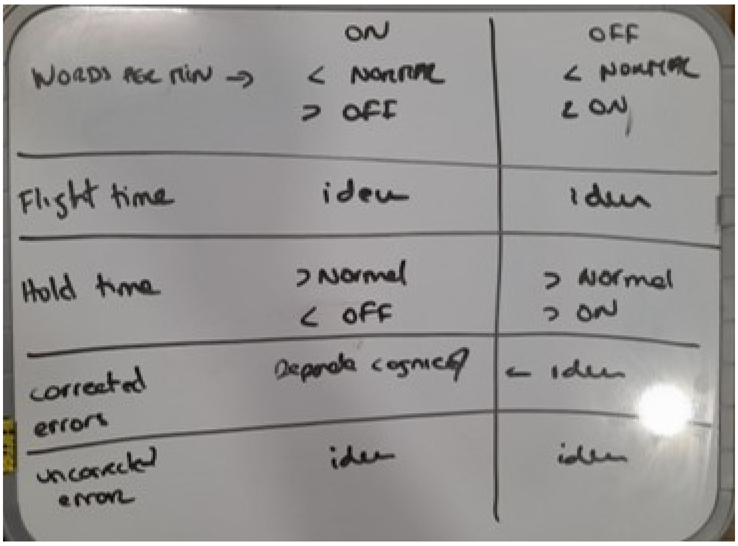

Figure 4.2: Example picture of the second exercise being carried out. Identification of expectations regarding the metrics.

Lastly, the participants were asked to draw on the white boards examples of how they would like to have the data represented, and the discussion followed, also taking into account three personas created: Persona 1 - 40s with mild symptoms, Persona 2 - 60s with mild symptoms, Persona 3 - 80s with aggravated symptoms.

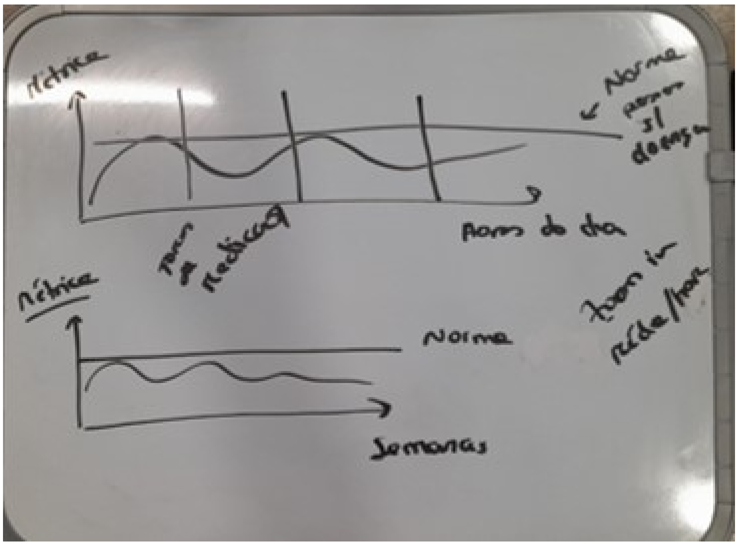

Figure 4.3: Example picture of the third exercise being carried out. Identification of data representation preferences.

4.4 Findings

4.4.1 Use of text-entry for neurodegenerative diseases

Even though the session with the clinicians was mostly focused on Parkinson's disease, as it is their expertise, with the symptoms similarity across different neurodegenerative diseases, the applicability of text-entry could be extended to the remaining diseases. This opinion is in line with previous research, as the systematic review showed, it is already possible to find work on Mild Cognitive Impairment and Multiple Sclerosis.

4.4.2 Potential text-entry for diagnosing and monitoring

When discussing the visuals and amount of data necessary to provide compelling insights, the clinicians mentioned it would be really interesting to have text-entry work towards diagnosing and monitoring. It was pointed out that depending on the patient's health condition, the amount of data needed would vary. For a diagnosis, it would be ideal to have as much previous data as possible to be able to understand the individual's own fluctuations and also, to explore the possibility of diagnosing from text entry behaviours alone.

As for monitoring, it was shared that it would be helpful, since sometimes clinicians are able to understand the patients' state by looking at them, however, some other times they need the history, and since it happens people have a weak support system and no info is provided from relatives observations, for example, it would be extremely beneficial to have a supporting tool for the patients monitoring. Regarding the amount of data, the minimum required would be a week prior to the patient's evaluation appointment, as this would already allow the clinician to understand the patient's state. The main issue pointed out by the clinicians was that, with the advancement of age and disease stage, patients tend to forget if they felt unwell and when it happened, at what time they took the medication, if they did it right or not, if they skipped a day, and if it was last week or three weeks ago. Hence, the more data the clinicians have available to them, the less they need to get directly from the patients.

The clinicians highlighted three exceptions for this "more from text-entry less from patients" rule. Firstly, as patients age, it is likely that they type less on their mobile devices, having less data to assess and impacting the clinicians' ability to extract any conclusions. Secondly, patients, independently of their age, can have agressive fluctuations on their cognition and become unable to use the device as expected. Lastly, as the disease progresses to later stages, patients become dependent on caregivers and stop interacting all together with devices, making the monitoring through text-entry obsolete.

4.4.3 Relationship between metrics and concepts of interest

From the clinicians point of view, all of the five initial metrics, words per minute, flight time, hold time, corrected error rate and uncorrected error rate, would seem to be impacted by the same concepts. Twelve concepts were highlighted by the clinicians. Starting with apraxia, which is the inability to perform a particular act, being learned or skilled, even though it can not be explained through elementary motor or sensory deficits or even through a language comprehension disorder [55]. The following identified concept was bradykinesia, characterised by slowness and difficulty in movement and loss of active motor ability [29]. The third concept identified was dyskinesia, which consists of involuntary movements associated with Levodopa therapy, which tends to be more predominant when patients are under the effect of medication [50]. The fourth concept was bradypsychia, which is basically the slowness of thought and mental activity [52]. The fifth concept was motor and non-motor fluctuations, which are variations in both motor and non-motor symptoms. These are invariably common in diseases like Parkinsons' and tend to happen in response to medication [21]. The sixth concept was cognitive defects which is identified by the decline of mental action and has multiple impacts on the patients' daily lives. As for the remaining six, the clinicians identified difficulty with dual-tasking (like thinking and writing at the same time), anxiety, tremor, tiredness, sleepiness and involuntary movements.

When moved on to the second set of nineteen metrics, the clinicians expressed that the same concepts they had just identified would also apply to the new metrics. In addition to this, it was possible to extract some conclusions regarding the clinician's expectations for the data.

According to them, it would be expected for a patient to have less words per minute than a healthy control (HC) during ON state but still higher than during OFF state, have lower flight time than a HC during ON state but still higher than during OFF state and have higher hold time than a HC during ON state but still lower than during OFF state. As for errors, clinicians would expect patients to have different corrected, uncorrected and total error rates depending on their cognitive abilities, since a person with higher cognitive defects may type more errors, regardless of ON and OFF states. The clinicians believe that for errors, it is mostly dependent on cognition rather than on motor capability, since a patient with bradypsychia, or with an attention deficit, may realise or not that the error is present and then may or not go back to correct it. However, for insertion error rate, they would expect for it to be higher than for an healthy control during ON state due to dyskinesia. It was also pointed out by the clinicians that omission and substitution error rates, key selected and motion info, would vastly depend on the patients' cognition abilities. However, if the patients' experience dyskinesia, this could also impact the results. Additionally, it would be expected for patients to have more time per word than a HC during ON state but still less than during OFF state, and a higher touch major / minor and offset than a HC during ON state due to dyskinesia.

4.4.4 Visualisation of text-entry metrics / practical potential

As for visualisations, the clinicians opted for line charts, where the metric being analysed would occupy the yy axis, and time would occupy the xx axis. It was identified that ideally, the time axis would be in hours with aggregated data from a selection of days right up until the evaluation day. It was also pointed out that it would be interesting to see events marked with vertical lines at specific times, such as Off Medication and Best On.

A discussion followed to understand how the clinicians needs would vary if the patients changed. The three personas were taken as examples and the clinicians believe that a younger person with a milder stage of the disease, like Persona 1, would still be able to report the past days, symptoms and fluctuations. Meaning that, it would be enough to see, instead of an hourly report, a week-by-week visual. For an older person with aggravated symptoms like Persona 3, without being totally incapacitated, the clinicians believe they have more difficulty reporting and therefore require more information from the data collected. It was pointed out that ideally, it would be possible to have the visuals with the option of drill-down through time, so whenever the necessity required, clinicians could see with more or less detail.

4.5 Summary

In an attempt to understand the clinicians' point of view on how text entry can contribute towards the diagnosis and monitoring of neurodegenerative diseases, CNS was approached to help. In a one hour session, the clinicians were shown five metrics and asked to correlate them with concepts of interest to them. Metrics shown: words per minute, flight time, hold time, corrected error rate, uncorrected error rate. Concepts correlated: apraxia, bradykinesia, dyskinesia, bradypsychia, motor and non-motor fluctuations, cognitive defects, difficulty with dual-tasking, anxiety, tremor, tiredness, sleepiness and involuntary movements. Following this exercise, the clinicians were asked to do the same correlation between said concepts of interest and another nineteen metrics collected by WildKey. The same relationship was identified to all metrics and concepts. Additionally, the clinicians expressed particular expectations for the data, regarding the patients ON and OFF states. Furthermore, the clinicians were asked to identify how they would like to see the data represented and finally, a discussion followed to understand how their visualisation needs would change according to the patient being evaluated.

Chapter 5

Discussion

Both the systematic review and work session with clinicians provide a better understanding of the work carried out so far on neurodegenerative diseases diagnosis and monitoring through text entry as well as the clinicians expectations for the metrics and future work.

Which Neurodegenerative Diseases have been Assessed/Monitored with Virtual Keyboards?
Seven of the nine articles found investigated Parkinson's disease with only one exploring Mild Cognitive Impairment and one Multiple Sclerosis. While all revealed promising results, there is an overall lack of research in this domain.

How was text-entry leveraged?
Prior work has focused on diagnosis, aiming to decrease the time period between the disease onset and disease diagnosis [26] [42]. One focused on exploring how text-entry could be used as a remote passively monitoring tool to assess motor decline, while one aimed to develop a remote PD test based on routine typing activities. Additionally, this review found that, according to the selected eligibility criteria, no articles focus on the monitoring of the disease, yet.

The split result between in-the-clinic and in-the-wild data collection highlights how research has looked at text-entry as a promising remote passive data collection method. While it can be used, and should be used in-the-clinic in early research, the use of text-entry seems to lie not in creating an additional clinical measure, but instead, focus on its ability to assess and monitor diseases in the wild. Additionally, from the studies reviewed, it appears that in-the-wild collection methods enabled researchers and practitioners to have larger cohorts.

What metrics are used?
The main metrics being collected are the hold time, flight time and pressure, showing a good correlation with symptomatic variation. While hold time and flight time are device independent, pressure values can vary from device to device [28] and their findings need to be subjected to additional work. Text-entry is a complex task that can produce a variety of different metrics that have yet to be explored within the context of neurodegenerative diseases (e.g. Error Rates), with the exception of Wang et al. 2021 [54] which focused on improving text-entry prediction algorithms for PD users. There is untapped potential in exploring if and how the different metrics associated with a text-entry session can be leveraged in the assessment and monitoring of neurodegenerative diseases. For PD the most common comparator for results validation is the UPDRS III test, that identifies Parkinson's disease.

5. DISCUSSION

This systematic review helped identify the lack of specification of the duration of the data collection. None of the present articles specified the timings when data was collected, being morning, afternoon or evening, and naturally there is no evidence of performance between time periods. No assessment was carried out regarding the effect of tiredness in the patient's results.

What were the main study outcomes?

As for prediction of test scores, one study found that there was a positive correlation between the performance of the dominant hand (for bradykinesia, hypokinesia and rigidity) and the UPDRS III item scores, holding true for both in-the-clinic and in-the-wild evaluations. Finally, it has also been shown that it is possible to improve patients' experience during their smartphone usage through text suggestion.

What are the clinicians beliefs on the potential of monitoring Parkinson's through text entry?

From the clinicians standpoint, text-entry can be a valuable tool towards the monitoring of neurodegenerative diseases. As the CNS clinicians expertise is more focused on Parkinson's disease, they were able to see clear connections between monitoring through text-entry and the patients' Parkinson's evolution. Particularly in situations where patients are debilitated and have a weak support system, they may not be able to remember exactly the progression of the disease in the weeks prior to the evaluation. Having the possibility of obtaining clearer information from an objective system would mitigate this difficulty and improve the clinicians ability to understand the patient's state.

As our study showed, clinicians believe that, from their experience with patients, text-entry metrics results would be affected by several different symptoms such as apraxia, bradykinesia, dyskinesia, bradypsychia and cognitive defects, among many others, demonstrating to be a promising area of research.

What are the clinicians expectations for the data?

Clinicians expect patients to have less words per minute than a HC, during ON state, but higher than during OFF state; have a lower flight time than a HC, during ON state, but higher than during OFF state; have a higher hold time than a HC, during ON state, but higher than during OFF state; have different error rates depending on the patients' cognitive abilities; have a higher insertion error rate than a HC, during ON state; have more time per word than a HC, during ON state but less than during OFF state; have a higher touch major / minor and offset than a HC during ON state. Clinicians expressed their expectation that dyskinesia in particular would affect the error rates and touch majors / minors.

What is the clinicians' experience regarding patients' evolutions? How does it typically impact their ability for text entry?

It was expressed by the clinicians that the evaluation of patients in different disease stages would require different amounts of data, but also that different disease stages impact the patients' abilities to cooperate with text-entry towards the disease monitoring. Meaning that, a patient with a milder disease stage, would still be able to recollect and report the previous days or even previous weeks, regarding feeling ill and taking medication, for example, making it less necessary to have huge amounts of data. On the other hand, patients with more advanced disease stages, but not yet later stages, typically have additional difficulty in remembering their progress, making the necessity of having other ways to collect the required information imperative.

What would be the ideal way to represent the data to aid in the diagnosis and monitoring of neurodegenerative diseases, namely Parkinson's disease?

The clinical team demonstrated that an ideal way for data representation, from their point of view, would be to have line charts, one for each metric, crossed against time in hours, with aggregated data from the previous weeks or months, right up to the evaluation day. Additionally, they expressed interest in seeing specific times of the day highlighted through vertical lines, for example, that would show when the patient was off medication, the patient's best ON, among other timestamps that may be curious to assess. The clinicians also expressed interest in having the possibility of drill-down through the visuals, across different time frames, depending on the patient being evaluated. Since a patient with a milder disease stage is expected to report easily, it would be enough for the clinicians to have the data aggregated by weeks, and if there was a particular week of interest, then they could drill-down and see the evolution by hours of that aggregated week. On the other hand, if the patient had a more advanced stage, the week-on-week representation may not be as crucial as the hour-on-hour. In a matter of efficiency, it would make sense to have one visual that would allow both representations.

Chapter 6

Conclusion

This dissertation showed the current state of the art through a systematic review as well as clinicians expectations through the work study. The systematic review shined a light directly at how neurodegenerative diseases have been researched in connection to text-entry. It provides the understanding that the majority of research carried out has been focused on Parkinson's disease. It was interesting to see that even though research identified four most common neurodegenerative diseases, Parkinson's, Alzheimer's, Huntington's and Amyotrophic Lateral Sclerosis [42], the articles found through the systematic review approached Parkinson's, Mild Cognitive Impairment connected to Alzheimer's, and Multiple Sclerosis. With only nine articles of interest, this shows that only a few articles have explored this matter and that they are all very scattered. As for the metrics the articles collected, the majority were focused on flight time, hold time and pressure. However, other studies have shown there are many more metrics that can be explored [47]. Regarding the time of collection, no standard duration for studies has been found and it seems that data collection time varies according to the number and duration of sessions. As for cohorts, it was interesting to see the clear age difference between patients and healthy controls, as it averages by ten years. As for the time with the disease itself, no standard way of reporting was found as some use disease onset and some opt for disease diagnosis. From the outcomes collected, it is possible to understand that the research carried out so far has been mostly focused on the collection and validation of metrics, and no research has included clinicians to provide their standpoint.

To contradict this tendency, a work session with clinicians was carried out to see whether it was possible to find additional connections. According to the clinicians, text-entry metrics do have the potential to be used for neurodegenerative diseases in general. When presented with the metrics found by Rodrigues et al. (2021) [47], clinicians were able to make connections with concepts of interest and twelve were found: apraxia, bradykinesia, dyskinesia, bradypsychia, motor and non-motor fluctuations, cognitive defects, difficulty with dual-tasking, anxiety, tremor, tiredness, sleepiness and involuntary movements. One of the main highlights from the session was that, contrary to what previous research has been focused on, what clinicians would expect to affect text-entry metrics in a more significant way would be cognitive declines rather than motor declines, and none of the articles found approach this issue. Additionally, clinicians not only provided a design suggestion for visuals and future dashboards, but also shared their necessity of data quantity to make sense to assess a patient, at least one week prior to the evaluation session, as well as expectations for data tendencies, such as a patient having less words per minute than a healthy control during ON state but still higher than during OFF state.

6.1 Benefits

As the research on this matter was so scattered, this dissertation provides a systemazation of said research. The review allows other researchers to understand the tendencies and gaps within the related work. In addition to this, technology is usually developed without the involvement of the main persons of interest, and so this thesis took a step further and the work was carried out together with clinicians to have their vision right from the beginning of the investigation, throwing hypotheses for future work.

6.2 Limitations

This thesis provides a review on previous work carried out up until April 2022, with research published only in three databases and from references of references, in english and after 2007. As the focus was text-entry through digital keyboards, other types of text-entry were not considered. Additionally, the work study was mainly focused on Parkinson's disease, as it was the clinicians expertise, possibly leaving behind different insights if there were clinicians with expertise in other neurodegenerative diseases. The visualisation outcomes were also very introductory, but should help future research with determining what to explore next.

6.3 Future work

From this exploratory study, it would now be interesting to investigate and assess how text-entry can actually help clinicians and patients with the diagnosis and monitoring of neurodegenerative diseases.

- Validation studies

Use the information collected from clinicians expectations and validate their hypotheses. Check whether in fact a patient would have less words per minute than a HC, during ON state, but higher than during OFF state, for example. Check with patients with different levels of cognitive impairment and see how error rates would vary. Additionally, it would be interesting to see data tendencies by groups. For example, taking a group of patients with similar clinical characteristics and seeing how their metrics vary, and the same for the other way, a group with similar metrics results and understanding the variations between that population.

- Data visualisation

As the aim is to help everyone involved with the disease, the data needs to reach the persons of interest. An easy way to do this is through visualisations accessible to them. However, it would be interesting to explore different visualisations for clinicians and for patients as their interests may be different.

- Longitudinal studies

Conduct longer studies to assess how text-entry metrics vary overtime for a group of people, together with their disease variations.

References

[1] Abdulaziz Alajaji, Walter Gerych, Luke Buquicchio, Kavin Chandrasekaran, Hamid Mansoor, Emmanuel Agu, and Elke A. Rundensteiner. "Smartphone Health Biomarkers: Positive Unlabeled Learning of In-the-Wild Contexts". In: *IEEE Pervasive Computing* 20.1 (2021), pp. 50–61. DOI: 10.1109/MPRV.2021.3051869.

[2] Teresa Arroyo-Gallego, Marıa J Ledesma-Carbayo, Ian Butterworth, Michele Matarazzo, Paloma Montero-Escribano, Verónica Puertas-Martın, Martha L Gray, Luca Giancardo, Álvaro Sánchez-Ferro, et al. "Detecting motor impairment in early Parkinson's disease via natural typing interaction with keyboards: validation of the neuroQWERTY approach in an uncontrolled at-home setting". In: *Journal of medical Internet research* 20.3 (2018), e9462. DOI: 10.2196/jmir.9462.

[3] Teresa Arroyo-Gallego, Marıa Jesus Ledesma-Carbayo, Alvaro Sánchez-Ferro, Ian Butterworth, Carlos S Mendoza, Michele Matarazzo, Paloma Montero, Roberto López-Blanco, Veronica Puertas-Martin, Rocio Trincado, et al. "Detection of motor impairment in Parkinson's disease via mobile touchscreen typing". In: *IEEE Transactions on Biomedical Engineering* 64.9 (2017), pp. 1994–2002. DOI: 10.1109/tbme.2017.2664802.

[4] Carlo Alberto Artusi, Murli Mishra, Patricia Latimer, Joaquin A Vizcarra, Leonardo Lopiano, Walter Maetzler, Aristide Merola, and Alberto J Espay. "Integration of technology-based outcome measures in clinical trials of Parkinson and other neurodegenerative diseases". In: *Parkinsonism & related disorders* 46 (2018), S53–S56. DOI: 10.1016/j.parkreldis.2017.07.022.

[5] Alzheimer's Association. "2019 Alzheimer's disease facts and figures". In: *Alzheimer's & dementia* 15.3 (2019), pp. 321–387. DOI: 10.1016/j.jalz.2019.01.010.

[6] GP Bates, R Dorsey, JF Gusella, MR Hayden, C Kay, BR Leavitt, et al. *Huntington disease. Nature reviews Disease primers. 2015; 1: 15005*. Tech. rep. 2015. DOI: 10.1038/nrdp.2015.527188817.

[7] Erika Bereczki, Rui M Branca, Paul T Francis, Joana B Pereira, Jean-Ha Baek, Tibor Hortobágyi, Bengt Winblad, Clive Ballard, Janne Lehtiö, and Dag Aarsland. "Synaptic markers of cognitive decline in neurodegenerative diseases: a proteomic approach". In: *Brain* 141.2 (2018), pp. 582–595. DOI: 10.1093/brain/awx352.

[8] Brian M Bot, Christine Suver, Elias Chaibub Neto, Michael Kellen, Arno Klein, Christopher Bare, Megan Doerr, Abhishek Pratap, John Wilbanks, E Dorsey, et al. "The mPower study, Parkinson disease mobile data collected using ResearchKit". In: *Scientific data* 3.1 (2016), pp. 1–9.

[9] Raquel Bouça-Machado, Constança Jalles, Daniela Guerreiro, Filipa Pona-Ferreira, Diogo Branco, Tiago Guerreiro, Ricardo Matias, and Joaquim J Ferreira. "Gait kinematic parameters in Parkinson's disease: a systematic review". In: *Journal of Parkinson's disease* 10.3 (2020), pp. 843–853.

REFERENCES

[10] Robert H Brown and Ammar Al-Chalabi. “Amyotrophic lateral sclerosis”. In: *New England Journal of Medicine* 377.2 (2017), pp. 162–172.

[11] Sonja von Campenhausen, Yaroslav Winter, Antonio Rodrigues e Silva, Christina Sampaio, Evzen Ruzicka, Paolo Barone, and et. al. Poewe Werner. “Costs of illness and care in Parkinson’s disease: an evaluation in six countries.” In: *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology* 21.2 (2011), pp. 180–191. DOI: 10.1016/j.euroneuro.2010.08.002.

[12] Grazia Cicirelli, Donato Impedovo, Vincenzo Dentamaro, Roberto Marani, Giuseppe Pirlo, and Tiziana R D’Orazio. “Human gait analysis in neurodegenerative diseases: a review”. In: *IEEE Journal of Biomedical and Health Informatics* 26.1 (2021), pp. 229–242. DOI: 10.1109/JBHI.2021.3092875.

[13] Lonneke ML De Lau and Monique MB Breteler. “Epidemiology of Parkinson’s disease”. In: *The Lancet Neurology* 5.6 (2006), pp. 525–535. DOI: 10.1016/S1474-4422(06)70471-9.

[14] Claudio De Stefano, Francesco Fontanella, Donato Impedovo, Giuseppe Pirlo, and Alessandra Scotto di Freca. “Handwriting analysis to support neurodegenerative diseases diagnosis: A review”. In: *Pattern Recognition Letters* 121 (2019), pp. 37–45. DOI: 10.1016/j.patrec.2018.05.013.

[15] Michael A DeTure and Dennis W Dickson. “The neuropathological diagnosis of Alzheimer’s disease”. In: *Molecular neurodegeneration* 14.1 (2019), pp. 1–18. DOI: 10.1186/s13024-019-0333-5.

[16] Neil Dhir, Mathias Edman, Álvaro Sanchez Ferro, Tom Stafford, and Colin J Bannard. “Identifying robust markers of Parkinson’s disease in typing behaviour using a CNN-LSTM network.” In: *CoNLL*. Association for Computational Linguistics. 2020, pp. 578–595. DOI: 10.18653/v1/2020.conll-1.47.

[17] Mariacruz L Dıaz. “Regenerative medicine: could Parkinson’s be the first neurodegenerative disease to be cured?” In: *Future Science OA* 5.9 (2019), FSO418. DOI: 10.2144/fsoa-2019-0035.

[18] Ezgi Dogan, Christian Sander, Xenija Wagner, Ulrich Hegerl, Elisabeth Kohls, et al. “Smartphone-based monitoring of objective and subjective data in affective disorders: where are we and where are we going? Systematic review”. In: *Journal of medical Internet research* 19.7 (2017), e7006. DOI: 10.2196/jmir.7006.

[19] Brittany N Dugger and Dennis W Dickson. “Pathology of neurodegenerative diseases”. In: *Cold Spring Harbor perspectives in biology* 9.7 (2017), a028035. DOI: 10.1101/cshperspect.a028035.

[20] Michael G Erkkinen, Mee-Ohk Kim, and Michael D Geschwind. “Clinical neurology and epidemiology of the major neurodegenerative diseases”. In: *Cold Spring Harbor perspectives in biology* 10.4 (2018), a033118. DOI: 10.1101/cshperspect.a033118.

[21] Susan H Fox and Anthony E Lang. “Motor and non-motor fluctuations”. In: *Handbook of clinical neurology*. Vol. 84. Elsevier, 2007, pp. 159–184. DOI: 10.1016/s0072-9752(07)84039-5.

[22] Luca Giancardo, Alvaro Sanchez-Ferro, Teresa Arroyo-Gallego, Ian Butterworth, Carlos S Mendoza, Paloma Montero, Michele Matarazzo, José A Obeso, Martha L Gray, and R Estépar. “Computer keyboard interaction as an indicator of early Parkinson’s disease”. In: *Scientific reports* 6.1 (2016), pp. 1–10. DOI: 10.1038/srep34468.

[23] Suneel Gupta. “Advances in levodopa therapy for Parkinson disease”. In: 86.14 (2016). DOI: 10.1212/WNL.0000000000002513.

[24] Orla Hardiman, Ammar Al-Chalabi, Adriano Chio, Emma M Corr, Giancarlo Logroscino, Wim Robberecht, Pamela J Shaw, Zachary Simmons, and Leonard H Van Den Berg. “Amyotrophic lateral sclerosis”. In: *Nature reviews Disease primers* 3.1 (2017), pp. 1–19. DOI: 10.1038/nrdp.2017.71.

[25] Dimitrios Iakovakis, K Chaudhuri, Lisa Klingelhoefer, Sevasti Bostantjopoulou, Zoe Katsarou, Dhaval Trivedi, Heinz Reichmann, Stelios Hadjidimitriou, Vasileios Charisis, and Leontios J Hadjileontiadis. “Screening of Parkinsonian subtle fine-motor impairment from touchscreen typing via deep learning”. In: *Scientific reports* 10.1 (2020), pp. 1–13. DOI: 10.1038/s41598-020-69369-1.

[26] Dimitrios Iakovakis, Stelios Hadjidimitriou, Vasileios Charisis, Sevasti Bostanjopoulou, Zoe Katsarou, Lisa Klingelhoefer, Simone Mayer, Heinz Reichmann, Sofia B Dias, José A Diniz, et al. “Early Parkinson’s disease detection via touchscreen typing analysis using convolutional neural networks”. In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE. 2019, pp. 3535–3538. DOI: 10.1109/embc.2019.8857211.

[27] Dimitrios Iakovakis, Stelios Hadjidimitriou, Vasileios Charisis, Sevasti Bostantjopoulou, Zoe Katsarou, Lisa Klingelhoefer, Heinz Reichmann, Sofia B Dias, José A Diniz, Dhaval Trivedi, et al. “Motor impairment estimates via touchscreen typing dynamics toward Parkinson’s disease detection from data harvested in-the-wild”. In: *Frontiers in ICT* 5 (2018), p. 28. DOI: 10.3389/fict.2018.00028.

[28] Dimitrios Iakovakis, Stelios Hadjidimitriou, Vasileios Charisis, Sevasti Bostantzopoulou, Zoe Katsarou, and Leontios J Hadjileontiadis. “Touchscreen typing-pattern analysis for detecting fine motor skills decline in early-stage Parkinson’s disease”. In: *Scientific reports* 8.1 (2018), pp. 1–13. DOI: 10.1038/s41598-018-25999-0.

[29] He Juanjuan, Wang Jianguo, Li Bochen, Yang Xianjun, et al. “An Automatic Detection Method for Bradykinesia in Parkinson’s Disease Based on Inertial Sensor”. In: *2020 IEEE 3rd International Conference on Electronics Technology (ICET)*. IEEE. 2020, pp. 166–169. DOI: 10.1109/ICET49382.2020.9119604.

[30] Münteha Nur Sonuç Karaboğa and Mustafa Kemal Sezgintürk. “Biosensor approaches on the diagnosis of neurodegenerative diseases: Sensing the past to the future”. In: *Journal of Pharmaceutical and Biomedical Analysis* 209 (2022), p. 114479. DOI: 10.1007/s00415-021-10831-z.

[31] Lampros C Kourtis, Oliver B Regele, Justin M Wright, and Graham B Jones. “Digital biomarkers for Alzheimer’s disease: the mobile/wearable devices opportunity”. In: *NPJ digital medicine* 2.1 (2019), pp. 1–9. DOI: 10.1038/s41746-019-0084-2.

[32] KH Lam, KA Meijer, FC Loonstra, EME Coerver, J Twose, E Redeman, B Moraal, F Barkhof, V de Groot, BMJ Uitdehaag, et al. “Real-world keystroke dynamics are a potentially valid biomarker for clinical disability in multiple sclerosis”. In: *Multiple Sclerosis Journal* 27.9 (2021), pp. 1421–1431. DOI: 10.1177/1352458520968797.

REFERENCES

[33] Bayard E Lyons, Daniel Austin, Adriana Seelye, Johanna Petersen, Jonathan Yeargers, Thomas Riley, Nicole Sharma, Nora Mattek, Katherine Wild, Hiroko Dodge, et al. "Pervasive computing technologies to continuously assess Alzheimer's disease progression and intervention efficacy". In: *Frontiers in aging neuroscience* 7 (2015), p. 102. DOI: 10.3389/fnagi.2015.00102.

[34] Walter Maetzler, Josefa Domingos, Karin Srulijes, Joaquim J Ferreira, and Bastiaan R Bloem. "Quantitative wearable sensors for objective assessment of Parkinson's disease". In: *Movement Disorders* 28.12 (2013), pp. 1628–1637.

[35] Pablo Martınez-Martın, Carmen Rodrıguez-Blázquez, Mario Alvarez, Tomoko Arakaki, Vıctor Campos Arillo, Pedro Chaná, William Fernández, Nélida Garretto, Juan Carlos Martınez-Castrillo, Mayela Rodrıguez-Violante, et al. "Parkinson's disease severity levels and MDS-Unified Parkinson's Disease Rating Scale". In: *Parkinsonism & related disorders* 21.1 (2015), pp. 50–54. DOI: 10.1016/j.parkreldis.2014.10.026.

[36] P McColgan and S J Tabrizi. "Huntington's disease: a clinical review". In: *European Journal of Neurology* 25.1 (2017), pp. 24–34. DOI: 10.1111/ene.13413.

[37] Seyed Zachariah Moradi, Faramarz Jalili, Negin Farhadian, Tanuj Joshi, Mingfu Wang, Liang Zou, Hui Cao, Mohammad Hosein Farzaei, and Jianbo Xiao. "Polyphenols and neurodegenerative diseases: focus on neuronal regeneration". In: *Critical reviews in food science and nutrition* 62.13 (2022), pp. 3421–3436. DOI: 10.1080/10408398.2020.1865870.

[38] Daniel L Murman. "Early treatment of Parkinson's disease: opportunities for managed care". In: *The American journal of managed care* 18.7 (2012), pp. 183–188.

[39] Anastasia Ntracha, Dimitrios Iakovakis, Stelios Hadjidimitriou, Vasileios S Charisis, Magda Tsolaki, and Leontios J Hadjileontiadis. "Detection of mild cognitive impairment through natural language and touchscreen typing processing". In: *Frontiers in Digital Health* 2 (2020), p. 567158. DOI: 10.3389/fdgth.2020.567158.

[40] Björn Oskarsson, Tania F Gendron, and Nathan P Staff. "Amyotrophic lateral sclerosis: an update for 2018". In: *Mayo Clinic Proceedings*. Vol. 93. 11. Elsevier. 2018, pp. 1617–1628. DOI: 10.1016/j.mayocp.2018.04.007.

[41] Matthew J Page, Joanne E McKenzie, Patrick M Bossuyt, Isabelle Boutron, Tammy C Hoffmann, Cynthia D Mulrow, Larissa Shamseer, Jennifer M Tetzlaff, Elie A Akl, Sue E Brennan, Roger Chou, Julie Glanville, Jeremy M Grimshaw, Asbjørn Hróbjartsson, Manoj M Lalu, Tianjing Li, Elizabeth W Loder, Evan Mayo-Wilson, Steve McDonald, Luke A McGuinness, Lesley A Stewart, James Thomas, Andrea C Tricco, Vivian A Welch, Penny Whiting, and David Moher. "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews". In: *BMJ* 372 (2021). DOI: 10.1136/bmj.n71.

[42] Alexandros Papadopoulos, Dimitrios Iakovakis, Lisa Klingelhoefer, Sevasti Bostantjopoulou, K Chaudhuri, Konstantinos Kyritsis, Stelios Hadjidimitriou, Vasileios Charisis, Leontios J Hadjileontiadis, and Anastasios Delopoulos. "Unobtrusive detection of Parkinson's disease from multi-modal and in-the-wild sensor data using deep learning techniques". In: *Scientific Reports* 10.1 (2020), pp. 1–13. DOI: 10.1038/s41598-020-78418-8.

[43] J. R. Playfer and J. V. Hindle. "Parkinson's disease in the older patient." In: *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology* (2008).

[44] M. Premkumar, S.R. Ashokkumar, V. Jeevanantham, S. Anu Pallavi, G. Mohanbabu, and R. Sathesh Raaj. "Design of cost-effective real time tremor alerting system for patients of neurodegenerative diseases". In: *Materials Today: Proceedings* 57 (2022), pp. 1989–1994. DOI: 10.1016/j.matpr.2021.10.187.

[45] Serge Przedborski, M Vila, and V Jackson-Lewis. "Series Introduction: Neurodegeneration: What is it and where are we?" In: *Journal of Clinical Investigation - J CLIN INVEST* 111 (Jan. 2003), pp. 3–10. DOI: 10.1172/JCI200317522.

[46] André Rodrigues, Hugo Nicolau, André Santos, Diogo Branco, Jay Rainey, David Verweij, Jan David Smeddinck, Kyle Montague, and Tiago Guerreiro. "Investigating the Tradeoffs of Everyday Text-Entry Collection Methods". In: *CHI Conference on Human Factors in Computing Systems*. 2022, pp. 1–15. DOI: 10.1145/3491102.3501908.

[47] André Rodrigues, André RB Santos, Kyle Montague, Hugo Nicolau, and Tiago Guerreiro. "WildKey: A Privacy-Aware Keyboard Toolkit for Data Collection In-The-Wild". In: *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*. 2021, pp. 542–545.

[48] Raymund AC Roos. "Huntington's disease: a clinical review". In: *Orphanet journal of rare diseases* 5.1 (2010), pp. 1–8. DOI: 10.1186/1750-1172-5-40.

[49] A Rosenblatt. "Neuropsychiatry of Huntington's disease". In: *Dialogues in Clinical Neuroscience* 9.2 (2007), pp. 191–197. DOI: 10.31887/DCNS.2007.9.2/arosenblatt.

[50] Albert Sama, Carlos Pérez-López, Jaume Romagosa, D Rodriguez-Martin, Andreu Catala, Joan Cabestany, David A Perez-Martinez, and Alejandro Rodrıguez-Molinero. "Dyskinesia and motor state detection in Parkinson's disease patients with a single movement sensor". In: *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 2012, pp. 1194–1197. DOI: 10.1109/EMBC.2012.6346150.

[51] Jonathan Synnott, Liming Chen, Chris D Nugent, and George Moore. "WiiPD—An approach for the objective home assessment of Parkinson's disease". In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 2011, pp. 2388–2391. DOI: 10.1109/IEMBS.2011.6090666.

[52] Eric Hartono Tedyanto and Estu Nila Widuri. "Bradypsychia and Urinary Incontinence Secondary to Frontal-Lobe High Grade Astrocytoma". In: *International Journal of Medical Reviews and Case Reports* 4.4 (2020), pp. 146–146. DOI: 10.5455/IJMRCR.Bradypsychia-Urinary-Incontinence-Secondary.

[53] Pichet Termsarasab. "Chorea". In: *CONTINUUM: Lifelong Learning in Neurology* 25.4 (2019), pp. 1001–1035. DOI: 10.1212/CON.0000000000000763.

[54] Yuntao Wang, Ao Yu, Xin Yi, Yuanwei Zhang, Ishan Chatterjee, Shwetak Patel, and Yuanchun Shi. "Facilitating text entry on smartphones with QWERTY keyboard for users with Parkinson's disease". In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021, pp. 1–12. DOI: 10.1145/3411764.3445352.

[55] Cindy Zadikoff and Anthony E Lang. "Apraxia in movement disorders". In: *Brain* 128.7 (2005), pp. 1480–1497. DOI: 10.1093/brain/awh560.

Appendices

Appendices

1. Systematic Review Protocol
2. Study with Clinicians Protocol

.1 Systematic Review Protocol

Smartphone text entry in assessing and monitoring neurodegenerative diseases: a systematic review protocol

Introduction

Previous studies regarding monitoring of neurodegenerative diseases have mostly included onsite appointments, which require both the patient’s self-report as well as collateral information, being provided either by a relative or a baseline test. This may present a problem, since onsite evaluations only capture a snapshot of that particular moment. Considering the reality of many outpatient clinics where patients can be evaluated seldomly throughout the year, the clinical impression of disease progression may be harder to attain (Lyons et al. 2015), which may impact therapeutic decisions, patient well-being and quality of life. Furthermore, patients’ self-report can be biased due to several reasons, including cognitive impairment and depressive symptoms or sleep disorders, making it difficult to ascertain fundamental details (Lyons et al. 2015, Martinez-Martin et al. 2014). Additionally, patient compliance may also impact these results.

There has been growing research on assessing symptoms and disease progression in a number of neurodegenerative diseases. Monitoring through typing analysis, attempting to mitigate some of the issues referred above, may allow healthcare teams to possess important data regarding symptom worsening and/or disease progression, collected as a free-living evaluation during the patient's usual activities of daily living. Therefore, the research question placed by this systematic review is: how is text entry being used towards assessing and monitoring neurodegenerative diseases?

Methods and analysis

Protocol and registration

We followed the reporting guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analysis for Protocols (PRISMA-P).

Prospero registration number **CRD42022293379.**

Study selection criteria

Inclusion

Articles will be included if they comprise studies which included patients with neurodegenerative diseases, and if published in peer-reviewed journals or conferences. As the main outcome of interest is to detail all research that has been carried out regarding monitoring of neurodegenerative diseases using touchscreen keyboards, articles with any type of metrics collection will also be considered, even if not directly connected to existing health-related scales.

Exclusion

Papers from 2007 onwards will be considered, since this date was established as the first time a full touchscreen, non-PDA, smartphone was released (Majumder and Deen, 2019). Additionally, articles published in languages other than English will also be excluded.

Search strategy

Electronic databases

The following databases were searched considering a 2007-2021 timeframe: PubMed, IEEE Explore, ACM Digital Library. An additional filter for English written articles was applied.

Search Terms

(Huntington OR "HD" OR Parkinson OR "PD" OR "Alzheimer" OR "AD" OR "Amyotrophic Lateral Sclerosis" OR "ALS" OR "Neurodegenerative" OR "Tremor" OR "Dementia" OR "Cognition") AND ("text-entry" OR "text-input" OR "text entry" OR "text input" OR "typing" OR "touch input") AND (smartphone OR touchscreen OR mobile OR tablet OR "Virtual Keyboard")

Manual searches

References included in originally considered studies will also be searched, if not found in the initial database search, in order to include other potential papers of interest.

Study selection

Any duplicated studies will be excluded by a reviewer (AM). Two reviewers (AM and AR) will independently assess 10% of the included articles to confirm inter-rater concordance. If no major issues arise, the remaining articles will be reviewed by one reviewer (AM). Any disagreements between the two reviewers will be resolved by a third reviewer (DB). Following PRISMA guidelines, a flow diagram will be made to represent the selection process.

Data extraction

Data extraction will be conducted by one author (AM) and disagreements will be resolved as described above. Non-English references will not be reviewed.

The data extracted will cover the following points:

- Study details: title, author, year
- Study design: duration, outcomes measured
- Participants' demographics: population size, inclusion and exclusion criteria
- Place of study: in-clinic and/or free-living
- Outcomes: metrics, compared to other baseline tests
- Results
- Conclusions

Additional Notes

Ethics and publication

This systematic review is exempt from ethics approval as the work is based on previously published work. It will also contribute towards a Master's thesis. The findings of the review will be published in a related peer-reviewed journal and presented at conferences.

References

Lyons, B. E., Austin, D., Seelye, A., Petersen, J., Yeargers, J., Riley, T., Sharma, N., Mattek, N., Wild, K., Dodge, H., & Kaye, J. A. (2015). Pervasive Computing Technologies to Continuously Assess Alzheimer's Disease Progression and Intervention Efficacy. *Frontiers in aging neuroscience*, *7*, 102. https://doi.org/10.3389/fnagi.2015.00102

Martínez-Martín, P., Rodríguez-Blázquez, C., Mario Alvarez, Arakaki, T., Arillo, V. C., Chaná, P., Fernández, W., Garretto, N., Martínez-Castrillo, J. C., Rodríguez-Violante, M., Serrano-Dueñas, M., Ballesteros, D., Rojo-Abuin, J. M., Chaudhuri, K. R., & Merello, M. (2015). Parkinson's disease severity levels and MDS-Unified Parkinson's Disease Rating Scale. *Parkinsonism & related disorders*, *21*(1), 50–54. https://doi.org/10.1016/j.parkreldis.2014.10.026

.2 Study with Clinicians Protocol

Co-Designing Text-Entry Endpoint Reports with Clinicians

Context and Goals

Previous studies regarding monitoring of neurodegenerative diseases have mostly included onsite appointments [1], which require both the patient's self-report as well as additional information, being provided either by a relative or a baseline test [1]. This may present a problem, since onsite evaluations only capture a snapshot of that particular moment [1]. Considering the reality of many outpatient clinics where patients can be evaluated seldomly throughout the year, the clinical impression of disease progression may be harder to attain [1] which may impact therapeutic decisions, patient well-being and quality of life. Furthermore, patients' self-report can be biassed due to several reasons, including cognitive impairment and depressive symptoms or sleep disorders, making it difficult to ascertain fundamental details [1][2].

There has been growing research on assessing symptoms and disease progression in a number of neurodegenerative diseases [4]. Monitoring through typing analysis, attempting to mitigate some of the issues referred above, may allow healthcare teams to possess important data regarding symptom worsening and/or disease progression, collected as a free-living evaluation during the patient's usual activities of daily living. In an attempt to improve the clinicians' ability to monitor neurodegenerative diseases, together with data collected from patients' text extry, this study aims to find a way to show the most relatable metrics collected through visualisations.

Participants

We will invite one group of three to five participants composed of clinical professionals that work with Parkinson's Disease patients. This includes neurologists and physiotherapists.

Procedure

This study will be composed of design sessions with the clinical professionals where participants will be engaged with the researchers in defining the most promising metrics and how their visualisations should be.
The study will start with a briefing, firstly with an introduction of the people involved, some words of appreciation for the participants time and an explanation of the purpose of this study.

> "Good morning, my name is (person), I am (description). These are (introduce the team). We sincerely appreciate your time to participate in this study which aims to help clinicians with an easier way to monitor neurodegenerative disease patients. For later reviewing purposes, this session will be recorded by the investigation team."

Secondly, an outline of the day will be made to make sure the participants are aware of the activities prepared.

> "Today we will start with a brief discussion focusing on your expectations regarding the data presented, to see what kind of connections you, the clinicians, would be expecting to see between the data and the patient's health reports. Afterwards we will do a practical session where you will create a draft of a dashboard and lastly we will present some visuals of two different patients to understand if the data makes sense to you."

Then a discussion session will follow, to understand any connections expected by the clinicians regarding the metrics collected and the patients' progression.

" Q1. What do clinicians believe to be the potential of monitoring Parkinson's through text entry?
Q2. What are the clinicians expectations?
Q3. What is the clinicians' experience regarding patients' evolutions? How does it typically impact their ability for text entry?"

The participants will be made aware of the metrics collected and will be asked to participate in an exercise to help determine the best way to see the data.

> "These are some of the metrics WildKey collects: number of words, number of errors corrected, number of errors remaining, flight time and hold time. The first three being the total numbers of each, the flight time being the sequence of values corresponding to the time between the release of two key taps, and the hold time being the sequence of values of time spent touching the screen in each touch. We would like for you to correlate these with concepts of interest, such as tiredness, cognition, speed and dexterity."

After this discussion, additional metrics will be shown to understand from the clinicians point of view, which might be interesting to have in a monitoring dashboard.

> "Besides the metrics presented there are more metrics available. Which of these metrics do you believe could be useful to see represented through visuals as well."

A slide presentation of all of the metrics will follow, with a quick explanation for each one.

The available metrics [3]:

- Speed
    - Words per Minute
    - Time per Word

- Errors
    - Corrected Error Rate - Of the characters erased, the percentage that was erroneous
    - Uncorrected Error Rate - Percentage of erroneous characters in the final transcribed sentence
    - Total Error Rate - Of the characters entered, the percentage that was erroneous, corrected, or not
    - Insertion Error Rate - Additional erroneous characters added (corrected and uncorrected)
    - Omission Error Rate - Corrected (characters that were missing at first but were backspaced and inserted) and Uncorrected Omitted characters in relation to the number of times the character was presented
    - Substitution Error Rate - Ratio of substitutions to intentions. We can provide both individually by letter and aggregated measures
    - Error Correction Attempts - Number of corrections sequences, meaning how many times the user started a sequence of correction actions
- Touch Dynamics
    - Flight Time - Sequence of values corresponding to the time between the release of two key taps
    - Hold Time - Sequence of values of time spent touching the screen in each touch
    - Touch Major/Minor - Sequence of values of the TouchMajor (length of the major axis of an ellipse that represents the touch area) and the TouchMinor (length of the major axis of an ellipse that represents the touch area)
    - Touch Offset - Sequence of values of the differences in key centroids and hitpoint deviations
    - Key Selected - The sequence of keypresses
    - Motion Info - Sequence of all the touch motions detected by Android and their timestamp
    - Timestamp
- Action and Character Counts
    - Action Count - Total number of actions performed
    - Correction Action Count - Total number of individual actions that corrected input
    - Entry Action Count - Total number of individual actions that produced an input
    - Number of Auto Corrects
    - Number of Changed Characters
    - Number of Selected Suggestions
    - Number of Written Characters
    - Number of Written Numbers
    - Number of Written Special Characters

After understanding which metrics the clinicians believe to be useful, they will be asked to create a prototype of the dashboard with all the visuals for the selected metrics.

> “Now that we know which visuals and metrics make sense, now we’d like to know what kind of report you would like to follow for patients’ monitoring.

> Here are the magnetic boards that you will use to make a sketch of a dashboard, selecting the visuals you think would work best for each metric. If there is any circumstance in which you believe another visual would be more suitable but is not available, please draw it with your pens."

Once the exercise is finished, the research team will debrief the invited members and say a few words of appreciation.

> "This research session is now over. We sincerely appreciate your time and effort towards the advancement of assessment and monitoring of neurodegenerative diseases through text entry."

Materials

For this session, the following materials will be required:

- Magnetic boards
- Pens
- Magnets with the possible visuals
- Computer / laptop
- Projector

References

1. Lyons, B. E., Austin, D., Seelye, A., Petersen, J., Yeargers, J., Riley, T., Sharma, N., Mattek, N., Wild, K., Dodge, H., & Kaye, J. A. (2015). Pervasive Computing Technologies to Continuously Assess Alzheimer's Disease Progression and Intervention Efficacy. *Frontiers in aging neuroscience*, *7*, 102. https://doi.org/10.3389/fnagi.2015.00102

2. Martínez-Martín, P., Rodríguez-Blázquez, C., Mario Alvarez, Arakaki, T., Arillo, V. C., Chaná, P., Fernández, W., Garretto, N., Martínez-Castrillo, J. C., Rodríguez-Violante, M., Serrano-Dueñas, M., Ballesteros, D., Rojo-Abuin, J. M., Chaudhuri, K. R., & Merello, M. (2015). Parkinson's disease severity levels and MDS-Unified Parkinson's Disease Rating Scale. *Parkinsonism & related disorders*, *21*(1), 50–54. https://doi.org/10.1016/j.parkreldis.2014.10.026
3. Rodrigues, A., Santos, A. R., Montague, K., Nicolau, H., & Guerreiro, T. (2021, September). WildKey: A Privacy-Aware Keyboard Toolkit for Data Collection In-The-Wild. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers* (pp. 542-545).
4. Wang, Y, Yu, A, Yi, X, Zhang, Y, Chatterjee, I, Patel, S, & Shi, Y (2021); Facilitating text entry on smartphones with QWERTY keyboard for users with Parkinson’s disease. Proceedings Of The 2021 CHI Conference On Human Factors In Computing Systems. doi: 10.1145/3411764.3445352