Graphical tasks to aid in diagnosing and monitoring Parkinson's disease

Smits, Esther

CHAPTER 7

DISCUSSION AND FUTURE DIRECTIONS

7.1. Introduction

The objective of the work described in this thesis was to design and evaluate measures for the quantitative assessment of motor symptoms of Parkinson's disease (PD). Currently, clinical assessment of PD mainly depends on the experience and interpretation of a clinician and diagnostic accuracy improves with increasing clinical experience[1]. Therefore, quantitative assessment of motor symptoms of PD might be useful to aid in diagnosing and monitoring PD. In this thesis we explored whether graphical tasks, such as handwriting and drawing, could be used to quantitatively assess motor symptoms of PD. Several important validation steps in diagnostic research, as described in section 1.3, were executed and their results described in the chapters of this thesis. The graphical tasks were performed and recorded with the DiPAR system, consisting of the DiPAR-pen, a digital tablet and operator computer, which was developed and evaluated in the European research project DiPAR (see Chapter 2 of this thesis). In the following paragraphs the results of the experimental studies are summarized and discussed, current and future clinical applications of the DiPAR system are discussed, and recommendations for future research are given.

7.2. Main results

The results described in Chapters 3, 4 and 5 show that standardized graphical tasks, as performed with the DiPAR system, could be used as an aid in the diagnostic process of PD. In Chapter 3, a proof of principle study is described that shows differences in performance between PD patients and healthy control (HC) participants. This is one of the first important steps in diagnostic research[2] to show that a new diagnostic test may be potentially useful clinically. As a second important step, we showed in Chapter 4 that the measures, which differed between PD patients and HC participants, were also highly reproducible in healthy adults. This means that measures are consistent over time, when no change in performance is expected. This result allows assuming that atypical performance of patients can be attributed to the disorder, rather than to a measurement error or learning effect[3]. Of course, for a test to be useful in the diagnostic process of PD, it is very important that it allows distinguishing between PD and other diseases with similar clinical symptoms. As early symptoms of PD could overlap with symptoms of other movement disorders (MD), the third step in assessing the diagnostic value of the new test was therefore to investigate differences between PD patients and patients with other

MD[2]. The results of this study are reported in Chapter 5 and confirm that a set of graphical tasks, as executed with the DiPAR-pen, can be used to assess differences in upper limb function between PD patients and patients with other tremor disorders. In addition to the evaluation of the diagnostic value of the graphical tasks, we also investigated the monitoring potential of such tasks in Chapter 6. We here showed that graphical tasks as performed with the DiPAR system provided valid measures to assess upper limb function in PD patients, and allowed detecting a response to dopaminergic medication.

Since the aim of this thesis was to develop and evaluate measures for the quantitative assessment of PD, we investigated whether specific graphical tasks and analyses provided clinically relevant information to assess the presence and severity of common motor symptoms of PD. In Chapter 3 we investigated measures related to two cardinal motor symptoms of PD, bradykinesia and tremor[4]. In addition, measures related to micrographia were studied, since micrographia is also a common symptom in PD[1,5]. These measures were subsequently used, improved and extended in the later experimental studies, described in Chapters 4–6.

Bradykinesia refers to slow movement and as a related measure we investigated movement time (MT) on the graphical tasks. According to the results described in Chapter 3, MT, especially on the writing tasks, showed significant differences between PD patients and HC participants. Writing tasks are more complex than simple tracing and drawing tasks and slowness of PD patients increases with increased movement complexity[6]. This does not necessarily mean that the simple tracing and drawing tasks are less useful than the writing tasks. The simpler tasks in fact also showed differences in MT between PD and HC participants and might be easier to perform correctly than the writing tasks, especially if such tasks are used for home-based monitoring. MT on the tracing and drawing tasks also showed high reproducibility in healthy adults, who performed the graphical tasks twice with one week in between, as described in Chapter 4. However, MT decreased on the second measurement day, which suggests a learning effect. This effect seems stronger in the easiest tracing task, the circle task, compared to the more complex tracing and drawing tasks, the spiral and zigzag tasks, which makes the circle task less reliable. In addition, MT on the simple tracing and drawing tasks, but also on the '*elelelel*' writing task, was significantly different between the four tremor groups, as described in Chapter 5. PD patients were slower, particularly in comparison with ET patients. Hence, MT on graphical tasks may be especially useful in distinguishing PD from ET patients, which can be clinically difficult[36]. Furthermore, Chapter 6 showed that graphical tasks were valid to assess bradykinesia in PD patients. In

particular, MT on the tracing and drawing tasks showed high correlations with an already validated measure for upper limb function, the Purdue pegboard test. Correlations for MT on the writing task with the Purdue pegboard test were weak and thus MT seems less valid. MT showed a response to dopaminergic medication in PD patients for all graphical tasks, as described in Chapter 6. According to their validity, reproducibility and ability to distinguish between PD and ET, the more complex tasks, like spiral and zigzag tracing and drawing, are suggested to be useful for assessing bradykinesia in PD patients. MT on writing as well as tracing and drawing tasks could probably be used to monitor treatment effects in PD patients.

Besides bradykinesia, we also investigated whether a second cardinal symptom of PD, *tremor*, can be assessed using a set of graphical tasks. We showed in Chapters 3, 5 and 6 that tremor can be detected during graphical tasks, as executed with the DiPAR-pen. Spectral analysis was used to investigate whether tremor was present during the tasks. The gyroscope sensors which were built in the second and third prototypes of the DiPAR-pen (see Chapter 2 for a description of the different prototypes) were especially useful to detect tremor. In line with previous studies[7,8], tremor frequency turned out not to be suitable to distinguish between different tremor disorders, because tremor frequencies overlapped too much between patient groups (Chapter 5). However, we did obtain some interesting results in Chapter 5, especially regarding the presence or absence of tremor during the graphical tasks (see Figure 5.3). We observed that all ET patients showed tremor on all tasks (posture task and the circle, spiral, and zigzag tracing tasks), whereas only a few EPT, FT and PD patients showed tremor on all four tasks (see Figure 5.3). Therefore the variable 'showing tremor on all four tasks', could be used to distinguish ET patients from the other tremor patient with high sensitivity (1.00). Additionally, in Chapter 5 we showed that tremor amplitude was low in EPT patients, in correspondence with previous literature[9,10]. In our study, tremor amplitude was significantly lower in the EPT group compared to the ET and FT groups. However, a few ET and FT patients also showed low tremor amplitude, so tremor amplitude could, in this study, not be used to distinguish EPT patients from the other tremor patients with high sensitivity and specificity. Furthermore, in Chapter 6 we investigated the effect of dopaminergic medication on tremor in PD patients performing the graphical tasks. In line with a study of Connoly et al.[11] we observed a decrease in tremor after taking dopaminergic medication in the PD patients who suffered from tremor during the graphical tasks.

Another common symptom of PD is *micrographia*. The DiPAR system can record performance during handwriting, and is therefore specifically useful

to quantify aspects of handwriting. Typically, handwriting difficulties in PD patients are characterized as micrographia, which refers to a reduction in writing size[5]. We have shown that the DiPAR system indeed allows to quantify writing size (Chapters 3–6). The '*elelelel*' task that we studied consisted of writing repeated patterns of the '*e*' and '*l*', which could automatically be recognized by a computer algorithm (developed by partner VTT in the DiPAR project, see Appendix 2). Subsequently, the size of the letters was calculated and used for statistical analysis. Such measures are expected to be useful in the diagnostic process of PD[5]. However, we observed that writing size was not always impaired in PD patients. While the PD patients in Chapter 3 wrote significantly smaller than HC participants, another group of PD patients showed a similar writing size compared to other tremor patients and HC participants in Chapter 5. This was in line with previous studies regarding micrographia, reporting that the prevalence of micrographia in cohorts of PD patients varies between 15% and 60%[12–16]. In addition, different prototypes of the DiPAR system, with different pens and tablets were used between the studies reported in Chapters 3 and 5, which may also have influenced the proportion of patients exhibiting micrographia.

The results of Chapter 5 suggest that writing size is not suited to aid in distinguishing different tremor disorders. Moreover, Chapter 6 showed that writing size measures only weakly correlated with the Purdue pegboard test making them less valid for assessing fine motor control and, corresponding to previous studies[15,17–19], they did not show a response to dopaminergic medication in PD patients. Yet, interestingly, most of the PD patients studied in Chapter 3 did not notice impairments of their handwriting in terms of writing size, even though we found a significant difference on this measure with HC. We therefore suggest that writing size could potentially be useful for the early detection of PD, but probably only in addition to other early symptoms of PD.

7.3. Towards clinical application

As shown in Chapters 3–6, graphical tasks provide measures to quantitatively assess upper limb function in PD patients. Such measures, obtained with the DiPAR system, could be used as an aid in diagnosing and monitoring PD. However, before a set of graphical tasks can actually be used in clinical settings, further validation studies are necessary and some improvements of the graphical tasks and the current prototype of the DiPAR system could be helpful. In this respect there are also some limitations of the experimental studies described in this thesis that need to be discussed. Firstly, the results of the

experimental studies described in this thesis cannot be generalized to the general population as the design was exploratory and only small groups of patients were included. This limits power and therefore the number of statistically significant results, but also reduces the probability that a significant result reflects a true effect[20]. Therefore, future adequately powered studies are needed to confirm and extend our results. However, for the DiPAR project, it was important to first show differences between small groups of PD patients, patients with other forms of tremor, and HC participants, before continuing development and production of the system on a larger scale. Additionally, the results of these exploratory studies were useful to investigate the usability of the tasks and to determine the average time needed to perform the tasks.

Secondly, as mentioned before, we found some conflicting results between studies. In Chapter 3 we found a difference in writing size between PD patients and HC participants. While we also studied handwriting in PD patients and HC participants in Chapter 5, we did not see a significant difference in writing size between these groups in that study. This contrast could be explained by the fact that a different prototype of the system was used for the studies in Chapters 3 (prototype V1) and 5 (prototype V2). None of the participants in our studies performed the graphical tasks twice with different prototypes of the system and therefore we were not able to investigate whether the differences in design indeed influenced performance on the graphical tasks. A cause for the observed differences could be that the surface of the tablet used for prototype V2 was smoother than the surface of the tablet used for prototype V1, which could make it easier for PD patients to perform smooth handwriting movements. Another difference between the studies described in Chapters 3 and 5 is that for the study described in Chapter 5 only tremor-dominant PD patients were included, while in Chapter 3 PD patients with severe tremor in the hands were excluded. Tremor-dominant PD is suggested to be a subtype of PD[21], and patients with this form of PD might therefore perform differently on graphical tasks than patients with other subtypes of PD. To our knowledge, no studies investigated the difference in presence of micrographia between different subtypes of PD. We expected that patients with tremor-dominant PD would show even more micrographia than patients with other forms of PD, since tremor in the hands supposedly influences activities performed with the hands. This was not the case, however (Chapter 5), which suggests that micrographia is actually less present in tremor-dominant PD patients. Nevertheless, we only studied a small group of tremor-dominant PD patients, so it would be interesting to further investigate the presence of micrographia in patients with different forms of PD in a larger study. An additional reason for the differences in findings between Chapters 3 and 5 might

be that PD patients in Chapter 5 were ON medication, while PD patients in Chapter 3 complied with overnight withdrawal of their medication. PD patients in Chapter 5 might therefore show similar performance to HC participants. On the other hand, in Chapter 6 we showed that micrographia measures (writing size) did not show a response to dopaminergic medication. The differences (medication status and subtype of PD) between the included PD patients studied in Chapter 3 and 5 can therefore not fully explain the conflicting results in Chapters 3 and 5. Hence, we propose that the differences between the hardware versions of the system largely explain the conflicting results. For future development of the system we thus suggest that the possible influence of different hardware on performance on graphical tasks should be taken into account.

7.4. Recommendations for future research

The experimental studies in this thesis indicate that the DiPAR system is reliable. It is valid to assess and monitor upper limb function in PD patients and it identifies differences between PD patients, patients with other tremor disorders and HC participants. All these steps are important in diagnostic research (Chapter 1). Nevertheless, future studies with larger groups of participants that are performed with the same final DiPAR system, to limit the influence of the hardware on performance, are needed to confirm these results.

Future studies could also investigate whether a set of graphical tasks, performed with the DiPAR-pen, may be used by non-specialists as triage test. Triage tests are used to increase the number of patients who will enter the clinical pathway, by picking up cases that otherwise could have been missed[3]. Triage tests need very high sensitivity to ensure that no cases are missed, because triage-negative persons will not be further tested[3]. Sensitivity of a clinical test refers to the ability of a test to correctly identify those patients with the disease[22]. For application as a triage test, future studies should investigate the sensitivity of the DiPAR system and it would be especially interesting to examine whether the DiPAR system allows to correctly identify patients with and without PD from a group of participants in whom it is clinically sensible to suspect PD[2]. This could involve patients with a family member with PD, which is one of the risk factors of PD, or patients with rapid eye movement sleep behavior disorder (RBD), which is one of the premotor markers of PD[23,24]. Several other risk factors, premotor markers and early motor symptoms have been described for PD by Lerche et al.[23] and are summarized in Figure 7.1. For example, age is a risk factor, because PD is more common in elderly people[25]. It

would thus be interesting to study and follow up elderly people with the DiPAR system and to assess its sensitivity for detecting PD in an early phase. With this application in mind, it would be interesting to add graphical tasks, performed with the DiPAR-pen, to large longitudinal cohort studies like 'Lifelines', in which individuals will be followed-up for 30 years[26]. Individuals older than 50 years could be asked to perform the DiPAR test battery in addition to the other Lifelines tests. Subsequently, these individuals could be followed-up to investigate who develops PD. However, such a longitudinal study takes many years to complete. As a more feasible alternative, one could first investigate whether patients with RBD or another premotor marker of PD can be distinguished from healthy persons (Figure 7.1), based on performance of graphical tasks using the DiPAR-pen.

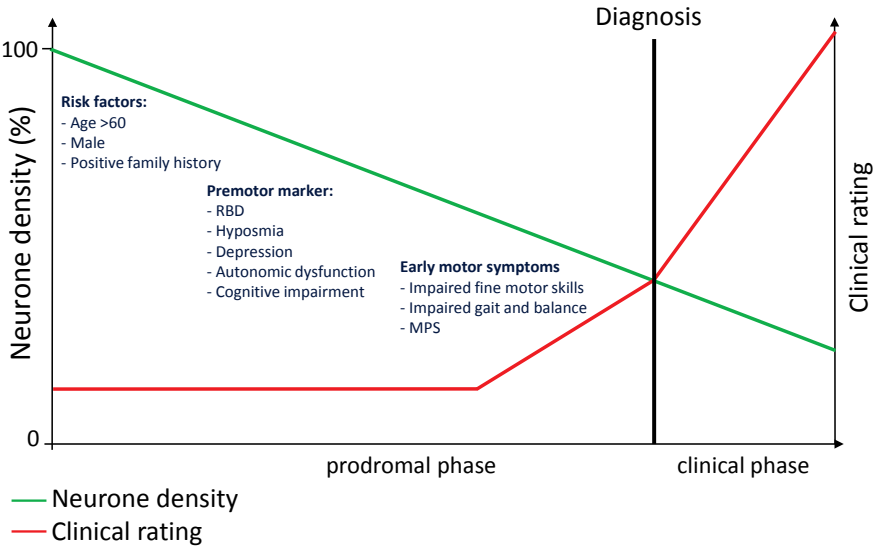

Figure 7.1. Adapted from Lerche et al.[23] Risk factors, premotor makers and early motor symptoms associated with loss of neurons (neurone density) in the prodromal phase of Parkinson's disease, based on the hypothesis of Braak et al.[27]. The decrease in neurone density is indicated in green, the increase in clinical rating scores in red. RBD=rapid eye movement sleep behaviour disorder; MPS =Mild parkinsonian signs[28].

Future longitudinal studies could also investigate how performance on the graphical tasks, performed with the DiPAR-pen, changes with time and which change is related to the development of PD and which change is related to healthy aging or other changes in lifestyle. This is an important step in evaluating whether graphical tasks performed with the DiPAR-pen could provide possible

biomarkers. Previously, it has been suggested that graphical tasks might provide biomarkers for PD[5]. Since the DiPAR system is non-invasive, easy to use and provides reproducible measures, it might provide suitable biomarkers, but more evaluation studies are necessary. For example, the values of a potential biomarker should be determined in the most healthy persons and the people who are most affected by a disease, in this case PD[29]. Chapter 4 of this thesis provides values for the most healthy persons and Chapters 3 and 5 provide values for PD patients. However, the measurements described in these chapters were not performed with the final, wireless prototype of the DiPAR system. Chapter 6 did provide values for PD patients who performed the tasks with the wireless DiPAR system, but these patients were generally only mildly affected. Our results should thus be confirmed and extended in future studies using the final product that Manus intends to launch, including larger groups of participants and more severely affected PD patients to determine standard values for the different groups of interest. As a start, we performed a prospective pilot study as part of the DiPAR project (the results of this study are not reported in this thesis), that included PD patients at different stages of the disease. However, these patients were all ON medication, while standard values of a biomarker should be determined for PD patients OFF medication.

It should be noted in this respect that although our results and previous studies suggest that kinematic analysis of handwriting and drawing tasks[5] indeed provides possible biomarkers for PD, it is unlikely that the assessment of (early) motor symptoms alone would be sufficient to diagnose PD in the prodromal phase (see Figure 7.1). An optimal early identification of PD will probably include non-motor measurements as well, such as a smell test, since hyposmia is one of the premotor markers of PD[23] (Figure 7.1). Such an approach is already common in clinical practice.

Besides being useful as an aid in the diagnosis of PD, our results suggest that graphical tasks performed with the DiPAR-pen, could be useful for monitoring purposes. Monitoring tools should accurately detect a change when it has occurred, which is the responsiveness of the tool[30,31]. In Chapter 6 we established a response to dopaminergic medication as a change in the performance on graphical tasks using the DiPAR system. However, we did not examine whether adequate doses of dopaminergic medications were given to the PD patients. Although it is common knowledge that dopaminergic medication is an efficacious treatment in PD, so that a response may be expected, a supramaximal dose should be given to PD patients to reliably assess responsiveness in future studies. Additionally, responsiveness of graphical tasks using the DiPAR-pen should be related to a change in a reference measure.

A change in the reference measure should be viewed as an accepted indication of a change, which is widely regarded by clinicians as meaningful and important, in the condition of a patient[32]. For PD, a change in UPDRS is widely regarded by clinicians as meaningful and important, if this change is large enough. In Chapter 6 we did assess UPDRS scores for PD patients OFF and ON medication, however, the changes in UPDRS scores between OFF and ON medication were not large enough to be used as a reference measure to investigate responsiveness of the graphical tasks, performed with the DiPAR-pen. For future studies, the minimal detectable change in UPDRS could be used as reference[33,34] to further determine the responsiveness of the graphical tasks performed with the DiPAR-pen. Furthermore, a monitoring tool should be able to detect a long-term change[35]. To investigate whether the DiPAR system can be used for long-term monitoring, PD patients should be followed for a longer time period in which they regularly perform the graphical tasks with the DiPAR-pen.

7.5. Conclusions

The results of the experimental studies described in this thesis show that graphical tasks, performed with the DiPAR-pen, can provide a quantitative assessment of upper limb function and motor symptoms in PD patients and patients with other MD. A set of standardized graphical tasks was shown to be useful as an aid in the diagnostic process of PD, to be valid to assess upper limb function in PD patients, and to be useful in monitoring short-term treatment effects. Bradykinesia and tremor measures were most helpful for both screening and monitoring of PD. Writing size measures might be useful for the early diagnosis of PD, but not for monitoring PD. According to their validity, reproducibility, and ease of use in the experimental studies, the circle, spiral, zigzag, '*elelelel*', and modified Fitts' tasks are recommended for use in future studies. The main advantage of using several graphical tasks, was that different motor symptoms of PD (bradykinesia, tremor, and micrographia) could be assessed simultaneously. In addition, such graphical tasks are non-invasive and easy to execute and the DiPAR system is portable, thus allowing home-based use without an examiner, which offers great opportunities for future clinical applications.

7.6. References

[1] Jankovic, J. Parkinson's disease: clinical features and diagnosis. *J. Neurol. Neurosurg. Psychiatry* **79,** 368–376 (2008).

[2] Sackett, D. & Haynes, R. Evidence base of clinical diagnosis - The architecture of diagnostic research. *Br. Med. J.* **322,** 1115–1117 (2001).

[3] Van den Bruel, A., Cleemput, I., Aertgeerts, B., Ramaekers, D. & Buntinx, F. The evaluation of diagnostic tests: evidence on technical and diagnostic accuracy, impact on patient outcome and cost-effectiveness is needed. *J. Clin. Epidemiol.* **60,** 1116–1122 (2007).

[4] Massano, J. & Bhatia, K. P. Clinical approach to Parkinson's disease: features, diagnosis, and principles of management. *Cold Spring Harb. Perspect. Med.* **2,** a008870 (2012).

[5] Letanneux, A., Danna, J., Velay, J.-L., Viallet, F. & Pinto, S. From micrographia to Parkinson's disease dysgraphia. *Mov. Disord.* **29,** 1467–1475 (2014).

[6] Moroney, R., Heida, C. & Geelen, J. Increased bradykinesia in Parkinson's disease with increased movement complexity: Elbow flexion - Extension movements. *J. Comput. Neurosci.* **25,** 501–519 (2008).

[7] Accardo, A., Chiap, A., Marino, S., Lanzafame, P. & Bramanti, P. Kinematic analysis of tremor - biomed 2010. *Biomed. Sci. Instrum.* **46,** 347–361 (2010).

[8] Elble, R. & Deuschl, G. Milestones in tremor research. *Mov. Disord.* **26,** 1096–1105 (2011).

[9] Buijink, a. W. G., Contarino, M. F., Koelman, J. H. T. M., Speelman, J. D. & van Rootselaar, a. F. How to Tackle Tremor – Systematic Review of the Literature and Diagnostic Work-Up. *Front. Neurol.* **3,** 1–12 (2012).

[10] Crawford, P. & Zimmerman, E. E. Differentiation and diagnosis of tremor. *Am. Fam. Physician* **83,** 697–702 (2011).

[11] Connolly, B. S. & Lang, A. E. Pharmacological treatment of Parkinson disease: a review. *JAMA* **311,** 1670–83 (2014).

[12] McLennan, J. E., Nakano, K., Tyler, H. R. & Schwab, R. S. Micrographia in Parkinson's disease. *J. Neurol. Sci.* **15,** 141–152 (1972).

[13] Contreras-Vidal, J. L., Teulings, H. L. & Stelmach, G. E. Micrographia in Parkinson's disease. *Neuroreport* **6,** 2089–2092 (1995).

[14] Kim, E. J., Lee, B. H., Park, K. C., Lee, W. Y. & Na, D. L. Micrographia on free writing versus copying tasks in idiopathic Parkinson's disease. *Park. Relat. Disord.* **11,** 57–63 (2005).

[15] Ling, H., Massey, L. A., Lees, A. J., Brown, P. & Day, B. L. Hypokinesia without decrement distinguishes progressive supranuclear palsy from Parkinson's disease. *Brain* **135,** 1141–1153 (2012).

[16] Wagle Shukla, A., Ounpraseuth, S., Okun, M. S., Gray, V., Schwankhaus, J., *et al.* Micrographia and related deficits in Parkinson's disease: a cross-sectional study. *BMJ Open* **2,** 1–7 (2012).

[17] Lange, K. W., Mecklinger, L., Walitza, S., Becker, G., Gerlach, M., *et al.* Brain dopamine and kinematics of graphomotor functions. *Hum. Mov. Sci.* **25,** 492–509 (2006).

[18] Tucha, O., Mecklinger, L., Thome, J., Reiter, A., Alders, G. L., *et al.* Kinematic analysis of dopaminergic effects on skilled handwriting movements in Parkinson's disease. *J. Neural Transm.* **113,** 609–623 (2006).

[19] Bidet-Ildei, C., Pollak, P., Kandel, S., Fraix, V. & Orliaguet, J. P. Handwriting in patients with Parkinson disease: Effect of l-dopa and stimulation of the sub-thalamic nucleus on motor anticipation. *Hum. Mov. Sci.* **30,** 783–791 (2011).

[20] Button, K. S., Ioannidis, J. P. a, Mokrysz, C., Nosek, B. a, Flint, J., *et al.* Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* **14,** 365–76 (2013).

[21] Thenganatt, M. A. & Jankovic, J. Parkinson disease subtypes. *JAMA Neurol.* **71,** 499–504 (2014).

[22] Lalkhen, A. G. & McCluskey, A. Clinical tests: sensitivity and specificity. *Contin. Educ. Anaesthesia, Crit. Care Pain* **8,** 221–223 (2008).

[23] Lerche, S., Hobert, M., Brockmann, K., Wurster, I., Gaenslen, A., *et al.* Mild Parkinsonian signs in the elderly - Is there an association with PD? Crossectional findings in 992 individuals. *PLoS One* **9,** 1–8 (2014).

[24] Miller, D. B. & O'Callaghan, J. P. Biomarkers of Parkinson's disease: Present and future. *Metabolism* **64,** S40–S46 (2015).

[25] Von Campenhausen, S., Bornschein, B., Wick, R., Bötzel, K., Sampaio, C., *et al.* Prevalence and incidence of Parkinson's disease in Europe. *Eur. Neuropsychopharmacol.* **15,** 473–490 (2005).

[26] Lifelines Research Office. LifeLines Cohort and Biobank. at <www.lifelines.net>

[27] Braak, H., Del Tredici, K., Rüb, U., De Vos, R. a I., Jansen Steur, E. N. H., *et al.* Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol. Aging* **24,** 197–211 (2003).

[28] Louis, E. D. & Bennett, D. a. Mild Parkinsonian signs: An overview of an emerging concept. *Mov. Disord.* **22,** 1681–1688 (2007).

[29] Maurits, N. M. Patiënten in getallen - Wiskunde in de klinische neurofysiologie (NL). 1–22 (2007). at <https://www.rug.nl/research/portal/files/2799159/oratie_z_dank.pdf>

[30] Roach, K. E. Measurement of Health Outcomes: Reliability, Validity and Responsiveness. *JPO J. Prosthetics Orthot.* **18,** P8–P12 (2006).

[31] Beaton, D. E., Bombardier, C., Katz, J. N. & Wright, J. G. A taxonomy for responsiveness. *J. Clin. Epidemiol.* **54,** 1204–1217 (2001).

[32] Husted, J. a, Cook, R. J., Farewell, V. T. & Gladman, D. D. Methods for assessing responsiveness: a critical review and recommendations. *J. Clin. Epidemiol.* **53,** 459–468 (2000).

[33] Heldman, D. A., Espay, A. J., LeWitt, P. A. & Giuffrida, J. P. Clinician versus machine: Reliability and responsiveness of motor endpoints in Parkinson's disease. *Park. Relat. Disord.* **20,** 590–595 (2014).

[34] Steffen, T. & Seney, M. Test-retest reliability and minimal detectable change on balance and ambulation tests, the 36-item short-form health survey, and the unified Parkinson disease rating scale in people with parkinsonism. *Phys. Ther.* **88,** 733–746 (2008).

[35] Bell, K. J. L., Glasziou, P. P., Hayen, A. & Irwig, L. Criteria for monitoring tests were described: Validity, responsiveness, detectability of long-term change, and practicality. *J. Clin. Epidemiol.* **67,** 152–159 (2014).

[36] Jain, S., Lo, S.E., Louis, E.D. Common misdiagnosis of a common neurological disorder. *Arch. Neurol.* **63**, 1100–1104 (2006).