



HÖGSKOLAN
DALARNA

Dalarna Doctoral Dissertations 9

Automating levodopa dosing schedules for Parkinson's disease

ILIASTHOMAS

Microdata Analysis
School of Technology and Business Studies
Dalarna University, Borlänge, Sweden
2019

Dissertation presented at Dalarna University to be publicly examined in sal Clas Ohlson, Borlänge, Friday, 5 April 2019 at 13:00 for the Degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Björn Eskifier (Friedrich-Alexander-Universität).

Abstract

Thomas, I. 2019. Automating levodopa dosing schedules for Parkinson's disease. *Dalarna Doctoral Dissertations* 9. Borlänge: Dalarna University. ISBN 978-91-85941-80-3.

Parkinson's disease (PD) is the second most common neurodegenerative disease. Levodopa is mainly used to manage the motor symptoms of PD. However, disease progression and long-term use of levodopa cause reduced medication efficacy and side effects. When that happens, precise individualized dosing schedules are required.

This doctoral thesis in the field of Micro-data analysis introduces an end-to-end solution for the automation of the pharmacological management of PD with levodopa, and offers some insight on levodopa pharmacodynamics. For that purpose, an algorithm that derives objective ratings for the patients' motor function through wearable sensors is introduced, a method to construct individual patient profiles is developed, and two dosing algorithms for oral and intestinal administration of levodopa are presented. Data from five different sources were used to develop the methods and evaluate the performance of the proposed algorithms.

The dose automation algorithms can work both with clinical and objective ratings (through wearable devices), and their application was evaluated against dosing adjustments from movement disorders experts. Both dosing algorithms showed promise and their dosing suggestions were similar to those of the clinicians.

The objective ratings algorithm had good test-retest reliability and its application during a clinical study was successful. Furthermore, the method of fitting individual patient models was robust and worked well with the objective ratings algorithm. Finally, a study was carried out that showed that about half the patients on levodopa treatment show reduced response during the afternoon hours, pointing to the need for more precise modelling of levodopa pharmacodynamics.

Keywords: Parkinson's disease, levodopa, symptom assessment, symptom management, dosing algorithms, wearable sensors, microtablets, continuous infusion

Ilias Thomas, School of Technology and Business Studies, Microdata Analysis,

© Ilias Thomas 2019

ISBN 978-91-85941-80-3

urn:nbn:se:du-29435 (<http://urn.kb.se/resolve?urn=urn:nbn:se:du-29435>)

*To all the people who made this thesis
possible and to those suffering from
Parkinson's disease.*

List of Papers

This thesis comprises the following papers, which are referred to in the text by their Roman numerals:

- I Thomas, I., Westin, J., Alam, M., Bergquist, F., Nyholm, D., Senek, M., & Memedi, M. (2018). A Treatment-Response Index From Wearable Sensors for Quantifying Parkinson's Disease Motor States. *IEEE Journal of Biomedical and Health Informatics*, 22(5), 1341-1349.
- II Thomas, I., Alam, M., Nyholm, D., Senek, M., & Westin, J. (2018). Individual dose-response models for levodopa infusion dose optimization. *International Journal of Medical Informatics*, 112, 137-142.
- III Thomas, I., Alam, M., Bergquist, F., Johansson, D., Memedi, M., Nyholm, D., & Westin, J. (2019). Sensor-based algorithmic dosing suggestions for oral administration of levodopa/carbidopa microtablets for Parkinson's disease: a first experience. *Journal of Neurology*, 1-8.
- IV Javed, F., Thomas, I., & Memedi, M. (2018, July). A comparison of feature selection methods when using motion sensors data: a case study in Parkinson's disease. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 5426-5429). IEEE.
- V Thomas, I., Memedi, M., Westin, J., & Nyholm, D. (2019). The effect of continuous levodopa treatment during the afternoon hours. *Acta Neurologica Scandinavica*, 139(1), 70-75.

Reprints were made with permission from the respective publishers.

Papers not included in the thesis:

- I. Thomas, I., Bergquist, F., Constantinescu, R., Nyholm, D., Senek, M., & Memedi, M. (2017). Using measurements from wearable sensors for automatic scoring of Parkinson's disease motor states: Results from 7 patients. *In Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE.*
- II. Johansson, D., Ericsson, A., Johansson, A., Medvedev, A., Nyholm, D., Ohlsson, F., ... , Thomas, I., Westin, J., & Bergquist, F. (2018). Individualization of levodopa treatment using a tablet dispenser and ambulatory accelerometry. *CNS Neuroscience & Therapeutics*, 24(5), 439-447.
- III. Crippa, A., Thomas, I., & Orsini, N. (2018). A pointwise approach to dose-response meta-analysis of aggregated data. *International Journal of Statistics in Medical Research*, 7(2), 25-32.

The following patent application has also been filed:

Sensidose AB, Authors: Ilias Thomas, Mevluding Memedi, and Jerker Westin (UK), "SYSTEMS FOR EVALUATING DOSAGE PARAMETERS", GB1708624.0, 2017.

Contents

Introduction.....	11
Dose titration and optimization.....	15
Research questions and contributions	16
Data description and processing	18
Dataset I	18
Dataset II	20
Dataset III.....	21
Dataset IV.....	22
Dataset V	23
Methodological framework.....	24
Individual model-fitting	24
Dosing algorithm for continuous infusion of levodopa.....	28
Dosing algorithm for oral administration of levodopa.....	29
Data analysis	33
Paper I	33
Paper II.....	34
Paper III.....	34
Paper IV	35
Paper V	35
Results.....	37
Paper I	37
Paper II.....	39
Paper III.....	39
Paper IV	42
Paper V.....	43
Conclusions.....	45
Discussion.....	47
Ethical considerations	49
Limitations	50

Future research.....	52
Acknowledgments.....	53
Περίληψη – Abstract (in Greek)	55
References.....	56

Abbreviations

CSAI	Continuous subcutaneous apomorphine infusion
CV	Cross-validation
DBS	Deep brain stimulation
ICC	Intra-class correlation
LID	Levodopa induced dyskinesia
LCIG	Levodopa-Carbidopa intestinal gel
LIO	Leave-one-out
MDS	Movement disorders society
PCA	Principal component analysis
PD	Parkinson's disease
PKG	Parkinson's Kinetigraph™
PKPD	Pharmacokinetic-Pharmacodynamics
PwP	People with Parkinson's disease
RMSE	Root mean squared error
SVM	Support vector machines
TRS	Treatment response scale
TRIS	Treatment response index from sensors
UPDRS	Unified Parkinson's disease rating scale

Introduction

2017 featured the 200-year mark since James Parkinson published the famous “An Essay on the Shaking Palsy”, where Parkinson’s disease (PD) was first noted and described (Parkinson, 2002). Having the disease leads to impairment of motor functions, as noted by Parkinson, and thus the disease is often classified as a movement disorder (Jankovic, 2008). Even though the incidence of the disease has not increased, the number of people with Parkinson’s disease (PwP) has increased considerably during the last decades, since the onset of the disease is often observed in people older than 65 years of age (Dorsey et al., 2007). The number of PwP is expected to rise even higher in the not-so-distant future since life expectancy has increased notably in most western countries, which will also be a burden on the healthcare system (Dorsey et al. 2013). Even though the disease was first noted over 200 years ago, the reasons for its onset are still unknown. As a result, no cure has been developed, and PwP have to live with the disease for the remainder of their lives.

The onset of PD reduces production of dopamine in the basal ganglia part of the brain, as neurons that produce dopamine begin to die. Reduced dopamine production leads to PD core symptoms, which are tremor, rigidity, postural instability and bradykinesia (Jankovic, 2008). Bradykinesia, as the Greek origin of the word implies, is slowness of movement. When the core symptoms are present then PwP exhibit Parkinsonism, which is also common in other diseases. The PwP are said to be in an OFF state when this behavior is present. Even though a cure is not available for PD there are treatment options. PwP on a treatment regimen will be referred to as patients.

The most common and effective pharmacological treatment for PD is levodopa (Connolly and Lang, 2014), which somewhat restores the dopamine level in the central nervous system, and as a result improves the patients’ motor functions to a normal level, referred to as ON. The patients’ response to levodopa is defined as “levodopa pharmacodynamics” and it is usually modelled through systems of differential equations. More details on pharmacodynamic models will be given in the “Individual model-fitting” chapter. Levodopa is usually used in combination with a stabilizer, such as carbidopa, and for the remaining of the thesis when referring to levodopa as monotherapy the use of a stabilizer is implied. In the early stages of the disease levodopa is delivered orally and a standard dose is sufficient treatment for the majority of the pa-

tients. Levodopa is used as monotherapy or in combination with other substances (e.g. dopamine agonists). However, as the disease progresses standard levodopa dosing is no longer sufficient. After five years with PD, 30-50% of patients start to have pharmacological side effects (Ahlskog and Manfred, 2001; Jankovic, 2005), such as motor fluctuations (going from OFF, to ON, to OFF, and so on), and also levodopa induced dyskinesia (LID) (Fabbrini, 2007). LID (manifestation of involuntary movements), as the name suggests, can be attributed to overmedication. As the disease progresses beyond the five-year mark, most patients will have those side effects. Personalized treatment is necessary to manage the disease manifestation severity in those cases and in general, after patient-specific dosing adjustments, most patients manage the disease adequately. Personalized treatment for oral medication refers to the time of day patients receive a dose and the amount of dose (Connolly and Lang, 2014). The number of doses per day vary considerably among patients (from three doses to nine, or more) and so does the amount of levodopa received in each dose. In general, the more the disease progresses the more doses are required to manage the symptoms and the side effects (Nyholm and Stepien, 2014). These drug therapies often have a narrow therapeutic window and produce significant side effects when dosing is non-optimal. The timing of doses, as well as the amount of the dose, is therefore critical to maintain drug levels within desired levels, and it is important that the administered doses are as accurate as possible to reduce the effects that can otherwise arise from over-, under-, or imprecise dosing. These treatments may involve a morning dose and then maintenance doses throughout the day. The need for a morning dose is also patient-specific, since there is a phenomenon called sleep benefit, and the morning dose intake can be postponed for a large percentage of patients (Merello, 1997).

It follows from the previous paragraph that patients increasingly require more and more doses as the disease progresses. Since it is infeasible or unrealistic to ask them to take a tablet every 30 minutes, more advanced therapies have been developed to allow for continuous infusion therapies (Timpka et al., 2016). The most prominent levodopa therapy available is the Levodopa-Carbidopa Infusion Gel (LCIG) device, which is installed in the abdomen area and provides continuous infusion of levodopa to the blood flow (Nyholm et al., 2005; Olanowor et al., 2006). Other surgical/invasive options are also available, such as continuous subcutaneous apomorphine infusion (CSAI), and deep brain stimulation therapy (DBS). Apomorphine is a dopamine-agonist and is also used to treat PD, to a lesser extent than levodopa (Timpka et al., 2016). In DBS, a device that provides electrical stimulation to the brain is installed (Volkman et al., 2013). CSAI and DBS are not further considered in this thesis study as the focus is mainly on levodopa use as monotherapy.

Before invasive therapies are considered, neurologists try to find the optimal oral administration dosing schemes for each patient. Since patients meet with the neurologists once or twice every year, the neurologist only gets a

snapshot of the patient's status and does not have a holistic understanding and representation of motor fluctuations and LID problems. The neurologist then relies on the patient's diaries, an assessment of the patient's appearance and physical characteristics, and discussions with the patient intended to reveal the patient's perceptions of response to the treatment and its side effects, which can be unreliable (Papapetropoulos, 2012). Those assessments are qualitative in nature and large disparities between expert assessments can be found in the literature (Heldman et al., 2011).

These issues might explain the ever-more expanding use of wearable sensors in clinical practice, both when remotely monitoring PD, allowing for a more comprehensive supervision of symptom manifestation, but also in an attempt to create objective measurements of the PD motor functions, to deal with the problem of between-expert variability. Many applications are found these days in the literature (Maetzler et al., 2016, Kubota et al., 2016), but it is worth noting that there are only two FDA-approved devices, the Parkinson's kinetigraph™ (PKG) (Griffiths et al., 2012), and the Kinesia™ systems (Giuffrida et al., 2009). Smartphone-based applications are also available, with UPDRScloud (Stamate et al., 2018) being the only one that has received CE marking. As in clinical practice, the monitoring devices/applications (approved or otherwise) derive scores for the PD status mainly in three states: OFF, ON and LID. Scores produced from wearables are often referred to as objective scores or ratings.

An interesting debate when considering wearable sensors or devices for monitoring Parkinson's disease symptoms concerns the most appropriate way to monitor. On the one hand, there is the position of the sensors on the body and if they should be on the hands, wrists, waist, etc. Also, there is the question if a combination of sensor placements would give the most information (Caramia et al., 2018). On the other hand, there is the question regarding passive or "instrumental" monitoring strategies. Instrumental monitoring, implies monitoring the patient when performing specific activities or instrumental tests, and passive monitoring when the patient is monitored in free-living activities. Another question of interest is the length of time for the monitoring and if it should be long-term or short-term. Many studies have been published in recent years debating these questions (Johansson, D., Malmgren, K., and Murphy, M. A., 2018) and the optimal use of sensors and how they could affect the patient's quality of life (Uem et al. 2016).

These objective scores facilitate in evaluating changes in the medication routines of the patient, and ratings are carried on motor function scales (Ebersbach et al., 2006; Ramaker et al., 2002). A variety of symptoms can be rated in these scales, such as the severity of the core symptoms (mostly tremor), or dyskinesia. In clinical practice, the patient performs instrumental tests and the treating personnel score the patient's motor functions on those scales. The most common instrumental tests are those used to assign scores on the Motor Examination part of the Movement Disorder Society - Unified Parkinson's

Disease Rating Scale (MDS-URDRS) (Goetz et al., 2008). Such instrumental tests can include: Arising from chair task, speech task, gait task, finger-tapping task, and pronation-supination movement of hands task. A commonly used scale in Sweden is the treatment response scale (TRS), a continuous scale that combines “OFF”, “ON”, and LID in the same range (Nyholm et al., 2005). The wearable devices can produce scores both during instrumental tests and during passive measurements, when patients perform free-living activities (Lopane et al., 2015; Griffiths et al., 2012; Giuffrida et al., 2009; Jane et al., 2016; Patel et al., 2009).

Dose titration and optimization

When individualizing the treatment (oral or continuous administration of levodopa), physicians make dosing adjustments to find a suitable dosing scheme for each patient. This process is often referred to as dose titration, and finding the optimal treatment might take multiple tries to achieve. During that period, the patient's response to different doses is observed and scored on the motor function states. The optimal scheme could either be the number and amount of doses for oral medication, or the infusion rate for continuous infusion treatment.

Commonly, dose titration decisions are based on the patients' opinions and physicians or specialized nurses past experiences when optimizing therapeutic effects. The dose titration process can take a considerable amount of time (up to ten days, in some cases) and usually involves over- and under-medication of the patients. During that period, patients usually experience potentially treatable symptoms and/or potentially unnecessary side effects, for a relatively long period of time. The wearable devices could assist the treating personnel during that process, monitoring the patient's motor functions (Mancini et al., 2012; Tsiouras et al., 2012; Dorsey et al., 2016; Espay et al., 2016).

However, when it comes to decision support for dosing adjustments through wearable sensors, expert systems are not currently available, thus, clinicians make dosing adjustments based on qualitative evaluations, even when objective scores are available. As a result, the optimal doses might take longer than necessary to achieve or might not be highly tailored to the patient's dosing needs. Furthermore, as the disease progresses, further adjustments are necessary and this process might need to be repeated many times during the lifetime of the patients. A potential hazard could be that patients who need regular dosing adjustment experience longer periods of non-optimal treatment, hampering their quality of life.

Research questions and contributions

The research goals of this thesis are to address some of the issues identified in the introduction. The research questions can be formulated as follows:

1. To what extent is it possible to map the motor functions of patients into a response scale that monitors the disease in a holistic way?
2. How can the information usually available in clinical practice be used to create individual patient profiles?
3. How can the individual patient profiles be used to make patient-specific dosing adjustments?
4. Do patients show lower response to the same pharmacological treatment during afternoon hours compared to the morning hours?

The thesis offers a solution for objective measurements, introducing a way to use instrumental tests to derive measurements on the TRS scale as a comprehensive way (holistic) of monitoring a patient's motor functions. Furthermore, a way to identify individual patient profiles for levodopa is introduced, and two dosing automation algorithms are described, one for oral administration of levodopa, and the other for LCIG. This unifies the thesis, as measurements on the TRS are imperative when building individual patient profiles, and the profiles are the input for the automation algorithms. The individual patient profiles will be referred to as individual models. These findings relate to research Questions 1, 2, and 3. Then, some insight on levodopa pharmacodynamics is offered, examining the patients' (non-) response to LCIG in the afternoon hours, which indicates that the dose-response models currently used might need some refinement, investigating research Questions 4. Research Question 4 is important in the context of the thesis study, since, without the use of appropriate dose-response models dosing automation cannot truly be optimal.

In the following chapters, a description of the five papers that comprise this thesis will be given. The primary contribution of the thesis is the introduction of two dosing algorithms that could potentially improve clinical practice. The first is related to the oral administration of levodopa and the second is related to LCIG infusion therapy. The algorithms are adjusted to clinical and objective ratings, demonstrating the feasibility of their use. A secondary contribution is the development of a way of deriving objective ratings, on a global scale that combines OFF, ON, and LID, which the algorithms can adapt to. Finally, an investigation of levodopa response is performed as a way of understanding the patients' dosing needs more extensively.

The thesis is in the field of Micro-data analysis, a normative science, meaning the objective and results were decision-support oriented, which was the overall aim of the thesis. Emphasis was placed on understanding the data, with the ambition of setting an early benchmark of the minimal requirements when making patient-specific dosing adjustments. This thesis makes a contribution in its field as it provides an algorithmic solution to individual patient needs through a data-driven decision-support process. Overall, the contribution of the thesis can be considered multidisciplinary as individualized medicine is an important part of the study, and health informatics methods were developed to address the research questions.

Data description and processing

Data from five different sources were collected and analyzed in the thesis. The data concerned both patients receiving oral levodopa treatment and patients on LCIG. The analysis of the first three datasets was the main focus of the thesis and two additional datasets were investigated for Paper IV.

These five datasets were part of the thesis because they contain the necessary information for exploring the research goals. All datasets were collected and analyzed after the research hypotheses were formulated, making the findings less susceptible to data snooping. Dataset I was collected to explore research Question 1, Dataset II is connected to research Questions 2 and 3, Dataset III to research Question 3, and Datasets II, IV and V to research Question 4. A detailed description of every dataset is given in the following paragraphs. Patients that were part of the studies will be referred to as subjects.

Dataset I

This dataset was used explicitly in Papers I and V, and implicitly in Paper III of the study. The data concerned subjects who enrolled in a levodopa challenge clinical study at Uppsala University Hospital during 2015. Nineteen subjects participated in that study, and the study design was approved by the Uppsala Ethics Committee. The subjects had given written informed consent for their participation according to the Helsinki declaration. The aims of the study were to explore the pharmacokinetics of a levodopa/carbidopa microtablet (Flexilev) formulation and to quantify the levodopa cycle objectively. For that reason, a levodopa challenge test was performed during the subjects' visit to the clinic. It was a multi-level study and data from different sources were collected. Blood samples were collected repeatedly during the levodopa cycle, which allowed for the analysis of the pharmacokinetics (movement of levodopa through the body compartments) of Flexilev (Senek et al., 2017). Furthermore, the subjects wore four six-degrees of freedom commercial sensors (Shimmer3) continuously, during the levodopa challenge, and performed a series of standardized instrumental tests repeatedly, at pre-defined time points, before, at the moment of, and following the single levodopa dose. The tests performed by the subjects were: speech test, finger-tapping, hand pronation-supination, foot-tapping, arising from a chair, gait sequence, spiral drawing, and eye-tracking movement. This series of tests will be referred to as a

“test session”. The aim of the tests was to collect data and develop algorithms that can objectively quantify PD motor functions in three states: OFF, ON, LID. Therefore, the subjects received 150% of their normal morning dose, to increase the probability of observing LID, and they also came into the clinic unmedicated, during their OFF states. During the instrumental tests, data from the four sensors were collected along with spiral drawing data and eye-tracking data. During every test session the subjects were video-recorded and the speech data were collected through the video recordings. The motor functions of the subjects in the videos were rated by three movement disorder experts. In total, there were 8-15 test sessions for each subject, and the recording were presented in a blinded order to the experts, who rated the subjects’ motor function independently of each other. The ratings were both on the TRS scale, and in sub-parts of the motor examination part of the MDS-UPDRS. Twenty two healthy controls wore the same sensors and had eight test sessions each, at pre-defined time points. I was not involved in the data collection process of this dataset, but a thorough description was required since it is the backbone of this thesis. The basic demographics of the participants in the study are seen in Table 1.

Table 1. *Population characteristics of the participants in study I. The standard deviation of the values is seen in the parentheses.*

	Sex	Mean age (years)	Mean height (m)	Mean weight (kg)	Mean years with the disease	Mean years on levodopa
Subjects	14 males 5 females	71.4 (6.3)	1.75 (0.09)	75.4 (11)	10 (6.8)	10 (6.8)
Healthy controls	16 males 6 females	64.2 (7.4)	1.75 (0.1)	83.6 (13.8)	-	-

The aim of the Uppsala study, among those mentioned above, was to investigate which combination of sensors and instrumental tests would provide the most information when quantifying patients’ motor functions. As a first step in this analysis, the signals from the wrist sensors were selected and processed, particularly, the signals during the pronation-supination hand rotation the subjects performed. The reason for this selection was, based on expert opinions, this movement was likely to contain most information. Raw signals were extracted from the sensors, and the 20-second hand rotation segments were identified visually, and cut from the signals for further analysis. For the 19 subjects who enrolled in the study, 229 20-second segments were extracted (about 12 test sessions per-subject, on average) and were further analyzed. The same 20-second segments were extracted for the healthy controls.

Dataset II

Similar to the Uppsala study, in Dataset I, a follow-up single-center longitudinal observational study was carried out in Sahlgrenska University Hospital at Gothenburg, Sweden, between August 2016 and February 2017. Thirty-one subjects enrolled, the study design was approved by the Regional Ethics Review Board at Gothenburg, and the participants had given written consent to the study, in accordance with the Helsinki declaration. The aim of this study was dose optimization for the oral administration of levodopa (Johansson et al., 2018), and it presented a few differences compared to the study described in Dataset I.

A major difference was that blood samples from the subjects were not collected during this study. Furthermore, the subjects received 120% of their normal morning dose making the observation of LID less likely. The distribution of the TRS observation among the two studies is shown in Figure 1. Another major difference was that during this study subjects had three visits (with a 2-week gap between each visit) to the clinic and between visits wore the PKG device for a six-day period in a free-living setting. During the second visit, the subjects repeated the same single-dose experiment as in the Uppsala study. In this case, the video recordings were rated by two movement disorder experts, who did the ratings together and reached a consensus rating for the motor functions of the subjects. Overall, the same data as in the Uppsala study were collected (minus the blood sample data, plus the PKG data) and these data were used in Paper III of the thesis study. No healthy individuals participated in this study. The demographic information of the subjects can be seen in Table 2. I was involved in the data collection process of this study from the first to the last subject visit.

Table 2. *Population characteristics of the participants in the single-dose experiments in Dataset II. The range of the values is seen in the parentheses.*

	Sex	Median age in years	Mean BMI as kg/m²	Median years from diagnosis	Median years with motor fluc- tuations
Subjects	15 males 10 females	68 (58-82)	25.2 (20.8- 35.4)	10 (4-30)	4 (1-20)

For this dataset, the same process as in Dataset I was followed, and from the 25 subjects who carried out the clinical experiment during their second visit to the clinic, 239 20-second segments of pronation-supination hand rotation were extracted. On average, the subjects in the Gothenburg study had a fewer series of tests (9.5 test sessions on average) compared to the Uppsala study. This was because the levodopa dose they received was a lower relative

amount, which in turn made the levodopa cycle shorter. The amount of dose the subjects received during the experiment was also known and used in the later analysis, and the dosing adjustments made by the supervising neurologist in the study, based on the PKG device, were used as the golden standard for oral administration of levodopa.

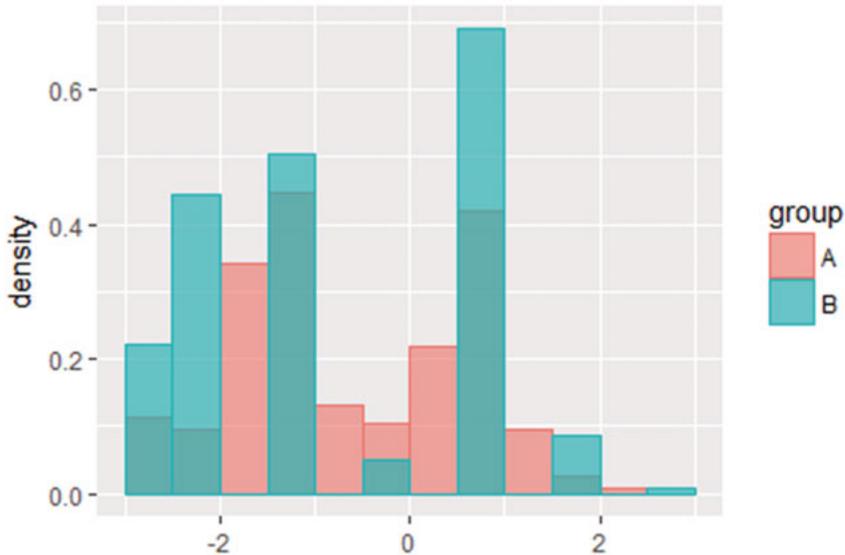


Figure 1. The histogram of the TRS values (x-axis) from Dataset I (A) and Dataset II (B). TRS values -3 and -2 in study II had about 3 times more observations than in study I.

Dataset III

This dataset was not collected during a clinical study, but the data were collected from Uppsala’s University Hospital patient database, retrospectively. An application was filed and approved by the Regional Ethical Review Board at Uppsala (application number 2016/361), and the participants had given written consent to retrospectively access their chart-data. The aim of this study was to investigate if automated dosing suggestions for continuous infusion are possible, and therefore LCIG data were collected. I wrote a draft for the ethical application that was filed, which was submitted by a movement disorder expert neurologist. These data provided information about LCIG dosing, and two search criteria were used: patients had to be diagnosed with idiopathic Parkinson’s disease, and had begun LCIG treatment in 2011–2015. There were 35 patients in the database that fulfilled those criteria, but four patients failed screening (two male, two female), since their files did not contain LCIG dosing information. Finally, data from 31 subjects were collected (nine female,

22 male). The data contained dose titration information and were in the form of scanned copies of the subjects charts. Most subjects spent a number of days in the clinic (on average, 5.7 days) during the dose titration process, and detailed dosing information (infusion rates, morning doses, and extra doses) was provided every day, along with ratings of the subjects' motor status on the TRS state. The ratings were carried out on an hourly basis from early in the morning until up to 22:00 hours by specialized nurses on the ward. These data were used in Papers II and IV of the thesis study. I collected the data from the Neurology Department database at Uppsala University Hospital, under the supervision of hospital personnel.

Processing of this dataset presented the most challenges as the records collected were mostly incomplete. The initial intention was to analyze all the available records for every subject, but that was not possible due to the high volume of missing values in the daily ratings. Consequently, it was decided that one day of data would be analyzed for each subject, the first day enough ratings were available (six was the minimum) and dosing information was known. This part of the dataset was used for Paper II. Data from the last day of admission for each subjects was also extracted and used in the analysis performed for the work in Paper II and IV. The last day of admission most likely contained the optimal dosing schedule for the subjects (since the titration process led to it) and was used as the gold standard when evaluating LCIG dosing schedules. Finally, data from 28 subjects were included in the analysis, as three subjects did not have dosing information for their discharge date. Demographic data for the subjects were not available, except for their sex.

Dataset IV

This dataset was from a previous study where subjects on LCIG treatment used an e-diary at home (Memedi et al., 2015). The aim of that study was to evaluate if PD treatment effects could be measured at home through self-entries in an e-diary. The subjects used the e-diary four times per day, at 08:00, 12:00, 16:00, and 20:00 hours, and self-scored their motor functions on the TRS (and other scales) while answering medication related questions during the process. The design of the study was longitudinal and the subjects used the e-diary for one week during nine different periods: baseline entry, and then 3, 6, 9, 12, 18, 24, 30, and 36 months after the baseline entry. There were two groups in this study, one naïve to LCIG, and the other non-naïve using the e-diary one additional time, three months before LCIG initiation. The dataset contained information on 65 subjects and these data were used for the study in Paper IV.

For this dataset a large number of records was available. The data from the period in which the subjects did not use LCIG (the naïve group before the baseline) was excluded from further analysis, as was one subject who only had

four observations. In total, 9386 observations from 58 subjects were included in the analysis and the observations were from nine observation periods for which the subjects used the e-diary.

Dataset V

The last dataset included in this thesis is from a 2003 study (Nyholm et al. 2003). Twelve subjects were enrolled in that study and the design was cross-over, where the subjects alternated between oral levodopa and LCIG treatment. The aim of the study was to calculate levodopa plasma concentrations for the two different treatments, and examine whether LCIG allows for more stable concentration throughout the day. The duration of the study was six weeks, and the motor functions of the subjects were observed for three days in total, during each treatment regimen. In addition to the assessments of the subjects' motor functions on the TRS scale, blood samples were also collected, on an hourly basis, from 08:00 hours to 17:00 hours. This dataset was also used for the study in Paper IV.

In comparison with Dataset IV, Dataset V had only a few observations for the 12 subjects. Three days of data were available for 10 subjects and two days of data were available for two subjects. For a third subject, the LCIG pump was disconnected on one of the trial days, and that day was excluded from the analysis. Each day had 10 hourly observations (from 08:00 hours to 17:00 hours) for both the TRS ratings and the levodopa plasma concentrations. There was one single exception where one of the days for a subject had only nine observations for the levodopa concentration, and the missing value was estimated based on the neighboring values (the value before and the value after).

Methodological framework

Before focusing on the data analysis, I believe that it is necessary to describe the novel methodologies developed to address the research questions. On the one hand, in order to individualize treatment, a technique was needed to identify what characterizes an individual dose-response. Thus, a method was devised to build individual dose-response models. On the other hand, related to the thesis goals, a method of using the individual models was necessary, and thus two distinct automation algorithms were developed, using the individual models as input. The methodological framework is described as follows: patient dose-response data \rightarrow individual models \rightarrow individual dosing suggestions through the dosing algorithms. In the following subchapters, the individual model building process and the two dosing algorithms are described. Parts of my licentiate thesis have been reused for this purpose (Thomas, 2017). Commonplace methods and algorithms will not be explained in a separate chapter, but will be briefly described in the data analysis section, if necessary.

Individual model-fitting

The research question that motivated this study was how individual patient information can be used to assess individual dose-response. In clinical practice, dosing information for patients is usually readily available and, provided monitoring is available, so are ratings of the patients' motor status on a treatment response scale (as described in Dataset III). The aim of this process was to explore how ratings on a specific scale and dosing information can be used together. The combination of these two elements should provide information about patients' specific needs, as they describe the patient-specific dose-response behaviors. This idea was inspired by the work of Jayachandran et al. (2014), where a method of predicting patients' specific needs for the treatment of leukemia was described, by individualizing dose-effect models.

In general, dose-response models are written as a system of mathematical equations and describe the response to a dose intake, for instance, a medicine intake, exposure to environmental factors, food consumption etc. In Jayachandran et al. (2014) the dose is radiation intake, and the response is how radiation treats leukemia. For medicine intake, it is common to employ pharmacokinetic-pharmacodynamics (PKPD) models to represent the dose-response function.

The dose-response model used in this thesis was specifically designed for levodopa intake. It is a two-compartment model, and the mathematical equations describe the movement (and concentration) of levodopa through the body compartments (Othman and Dutta, 2014; Simon et al., 2016; Westin et al., 2011), namely from levodopa intake until finally reaching an effect compartment. Thus, the concentration of levodopa in the effect compartment manifests as a value on an effect scale. The selected PKPD model estimates the effect on the TRS scale (Westin et al., 2011), which is used throughout the thesis. The reason for this choice was that it made intuitive but also practical sense to select a scale where the patients' motor functions are modelled in a continuous fashion, and it is also a scale used broadly in Sweden. The system of equations for the dose-response model used can be seen in Equations 1-5.

$$\frac{da_0}{dt} = \text{Inf} - k_a * a_0 \quad (1)$$

$$\frac{da_1}{dt} = \text{BIO} * k_a * a_0 - \left(\frac{Q+\text{CL}}{V_1}\right) * a_1 + \left(\frac{Q}{V_2}\right) * a_2 + R_{\text{syn}} \quad (2)$$

$$\frac{da_2}{dt} = \left(\frac{Q}{V_1}\right) * a_1 - \left(\frac{Q}{V_2}\right) * a_2 \quad (3)$$

$$\frac{dc_e}{dt} = \text{kEO} * \left(\frac{a_1}{V_1} - c_e\right) \quad (4)$$

$$E = \text{BASE} + \frac{E_{\text{max}} * c_e^\gamma}{c_e^\gamma + EC50^\gamma} \quad (5)$$

This model can be described as a two-compartment model, where dose intake takes place in Equation 1, and after movement through the body compartments, it reaches the effect compartment in Equation 5, and effect is manifested in response to levodopa concentration. The parameter and variable description of the PKPD model (Equations 1—5) is given in Table 3.

Table 3. *Description of Equations 1-5*

Name	Explanation	Variable / parameter
Inf	Intestinal levodopa infusion rate (mg/min)	-
a0	Amount in first compartment (mg)	Variable
a1	Amount in second compartment (mg)	Variable
a2	Amount in third compartment (mg)	Variable
k_a	Absorption rate (1/min)	parameter
TABS	1/k _a , absorption time constant (min)	parameter
kEO	Effect rate (1/min)	parameter
TKEO	1/kEO, effect time constant (min)	parameter
BIO	Bioavailability	parameter
Q	Intercompartmental clearance (L/min)	parameter
V1	Volume in first compartment (L)	parameter
V2	Volume in second compartment (L)	parameter
CL	Clearance rate (L/min)	parameter
Rsyn	Endogenous levodopa synthesis rate (mg/min)	parameter
Ce	Concentration in the effect compartment (mg/L)	Variable
EC50	Concentration at 50% effect (mg/L)	parameter
Gamma	Sigmoidicity factor	parameter
BASE	Baseline effect	parameter
Emax	Change from baseline effect	parameter
E	Effect ranging from -3 to +3	-

In Table 3, the model contains both parameters and variables. In order to fit individual models, the strategy was to estimate a few patient-specific parameter values. The patient-specific parameters were selected through a sensitivity analysis. The parameters that were not patient-specific were fixed to the population mean values, as in Westin et al. (2011). The sensitivity analysis was carried out to evaluate which parameter variations most influence the effect value. In total, 4 parameters were chosen to be individualized (together with an added secondary parameter), and the rest were set to population mean values. To individualize the 5 parameters, mathematical optimization was applied, by minimizing an objective function through least squares minimization. The objective function can be seen in Equation 6, where the predicted effect is the value output from Equation 5, and the TRS values are the scores assigned on the TRS scale (by experts or sensor scores). A schematic illustration of the individual model-fitting process can be seen in Figure 2.

$$Objective_function = \sum_{k=1}^n (Predicted_effect_k - TRS_values_k)^2 \quad (6)$$

To evaluate the ability of the method to convert observed dose-response behaviors to individual dose-response patient profiles, a simulation study of continuous infusion was conducted. This process was carried out to assess if this method had enough flexibility to fit individual models. For that experiment 100 profiles were generated, based on the population values of Westin et al.

(2011) through a covariance matrix (individual parameter values for all profiles in that study were available). For each of the 100 profiles, an infusion morning rate and maintenance rate was provided and, with a sampling rate of 60 minutes, 12 observations were collected from each simulation (without sampling or added error), with the values of the profile's response on the TRS scale. For all profiles, the algorithm managed to fit individual models, with the objective function value of virtually zero. In other words, the mean residual value from the model-fitting process was zero.

In that step, multiple optimization algorithms were evaluated, all giving similar results. Based on those results, the model individualization strategy was assessed as sufficient and was succeeded by application of this methodology on real-life data. Of course, in real-life data the objective function value will hardly ever be zero, since the TRS values will likely contain rating error (both from experts and from sensor predictions), or patients might experience motor fluctuations, which cannot be accommodated by the current PKPD models.

After developing the model-fitting algorithm, there was still a gap in knowledge about how the individual models can be used to make dosing suggestions. For that reason, two dose automation algorithms were developed, taking individual models as input, one for continuous infusion of levodopa, and one for oral administration of levodopa. Both algorithms were applied in real-life data and the suggestions they produced were evaluated against dosing decisions taken by movement disorder experts, used in this thesis as the gold standard.

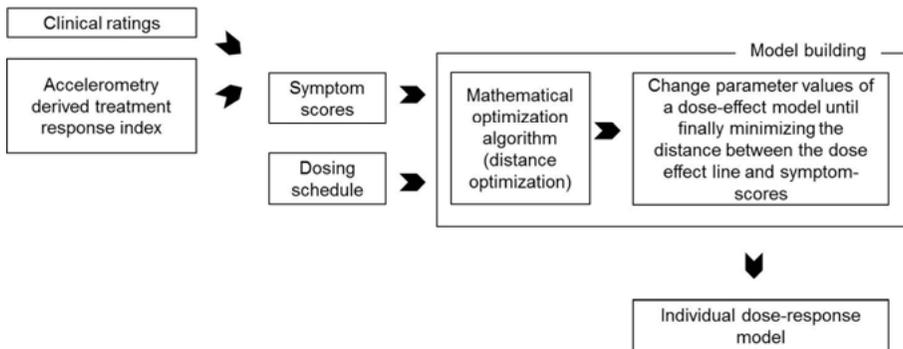


Figure 2. Schematic illustration of the individual model-fitting process.

Dosing algorithm for continuous infusion of levodopa

Continuous infusion optimization is not something new as control theory has been extensively used in medical applications, when there is a known target value for a physiological measurement. The most notable application of control theory can be found in the treatment of diabetes (Leelarathna et al., 2014), where an infusion pump supplies the patients with insulin, and the amount of insulin delivered is determined by the concentration of glucose in the blood flow. The application can be described as a closed-loop process, as the system adjusts the dosing based on the output measurements. The target in that case is the glucose levels in the blood.

Another noteworthy application of continuous infusion optimization is the anesthesiology field, where both closed and open loop feedback systems are used in practice (Struys et al., 2016). In this case, the target is the desirable clinical state of the patients.

In this study the target value for the optimization procedure is a value selected on the TRS scale as the optimal state of the patients. This target value would be zero, as it is the ON state of the patients. As far as continuous infusion is concerned, treating personnel can find the optimal treatment quite accurately during the titration period. That is because the therapeutic decisions mainly concern the correct choice of the constant rate, a task of low complexity. The challenge in this case is not in finding the optimal doses, but doing so in a timely fashion. Dose-titration takes days, and patients usually are admitted to the hospital for this, even though it has been also applied in remote fashion (Willows et al., 2016). Having algorithmic assistance that could help in reducing the dose-titration period has motivated the construction of the algorithm.

For the design of the continuous infusion algorithm, the LCIG pump specifications were taken into consideration. Through that device a morning dose is given at the beginning of the day, and after the dose is supplied a continuous infusion rate is initialized. The morning dose is high, and is aimed at bringing the patients to an ON state quickly at the beginning of the day. The continuous infusion rate (maintenance dose) is lower and assists with maintaining the patient on the ON state, since it sustains the levodopa plasma concentration stable.

The limitation of this device, compared to the device used for diabetes treatment, is that sampling of levodopa plasma concentration is not possible, since, to calculate levodopa plasma concentrations, blood samples need to be examined at a laboratory, whereas glucose levels are easily found in the blood flow. That means that the motor state of the patient needs to be assessed by an expert (or with wearable sensors), and a specific morning dose and maintenance dose need to be manually programmed into the device.

As seen in Figure 3, the optimization strategy entails minimizing an area (dark shade in the figure), which, in that example consists of three stages: before reaching the ON state; with a slight overshoot; and when the TRS values

go into the negative again. This is a typical representation of the optimal behavior of a patient, even though there can be exceptions. Mathematical optimization is used to minimize those areas, or, more precisely, to minimize the distance between the dose-effect curve and the target.

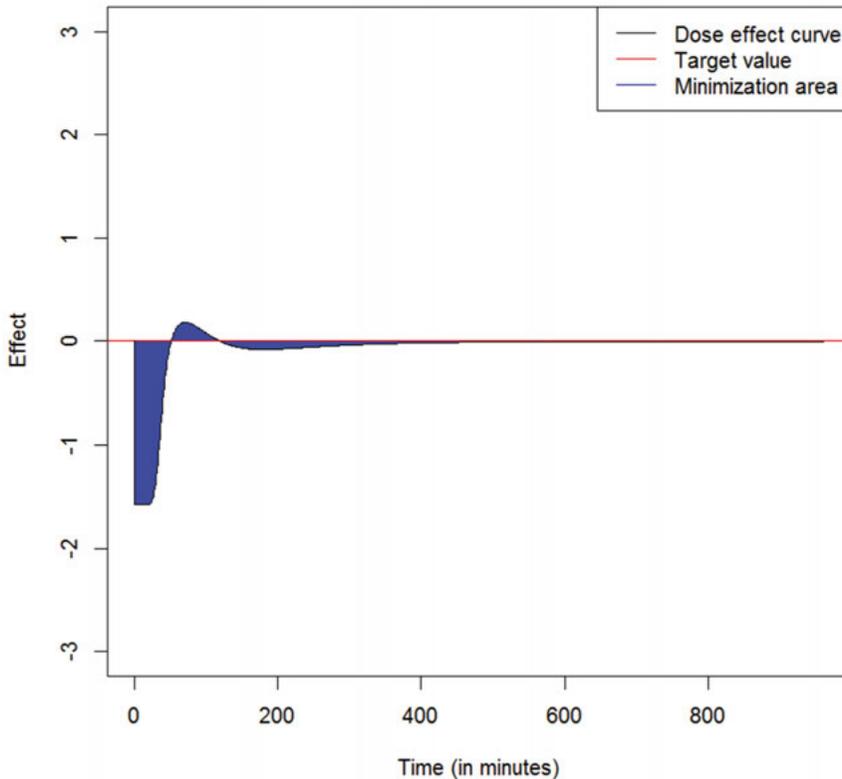


Figure 3. Optimization strategy for deriving optimal morning dose and infusion rate. The morning dose is large enough for the effect to reach the target value as soon as possible, and, at the same time, it is small enough for the overshooting to cause minimal side-effects. The algorithm adjusts the infusion rate and ensures that once the effect line reaches the target value again, it remains there throughout the day (960 minutes here – 16-hour day).

Dosing algorithm for oral administration of levodopa

In Paper III, the research question moves focus from continuous infusion to oral medication of levodopa, particularly at the advanced treatment stage. In contrast with the continuous infusion, the complexity of this problem, even for experts, makes the determination of an optimal dosing scheme (dosing interval and amount of dose) a challenging task (Nyholm and Stepien, 2014).

In this case, the construction of a dosing algorithm is not as straightforward, and multiple aspects need to be considered. To begin with, there is the fundamental understanding that this way of medicating is not optimal for treatment efficacy. A naïve but instinctively relevant question would be why all patients do not have an infusion pump or DBS (to receive improved treatment), and why treating personnel must bother with tailoring individual doses when superior options are available. These questions have a multilevel answer which relates to the cost of treatment: an infusion pump is by far the most expensive (Valledeoriola et al., 2013); potential side-effects must be considered. For instance, DBS has the most severe side-effects (Fenoy and Simpson, 2014); even though there is evidence to suggest that continuous infusion can cause peripheral neuropathy (Mancini et al., 2014), convenience for the patient must be considered (for instance, the poor quality of life for the patient with a device attached to the body); and last, but most definitely not least, the potential improvement of treatment for the patient. A comprehensive evaluation of these factors is performed to determine if the potential treatment improvement outweighs the combined adversity risks of the above factors.

Thus, it is now possible to examine the dosing algorithm for oral administration more closely. Patients experience a wearing-off effect from one dose to the next, and most patients will have dyskinesia manifestation during the day. When designing the algorithm, adverse behavior was a given, and the system is built to minimize these adverse effects. The question is how one finds a dose that is high enough for the ON state to be reached but, at the same time, it is low enough for overmedication to be avoided.

This question was answered by setting a target interval for the TRS, and keeping the expected dose-response effect within that interval as long as possible. In Figure 4, the dosing algorithm design is graphically depicted. For a fixed dosing interval (time between two doses), a morning dose is given, with subsequent maintenance doses throughout the day. The estimated effect is kept within a target interval range and an additional safety is set, with a threshold value to prohibit the fluctuation from becoming so large that the patient's condition swings back and forth in different motor function states. The imposition of a threshold value, which the dose-effect curve could not cross is called the fluctuation criterion.

Optimization strategy

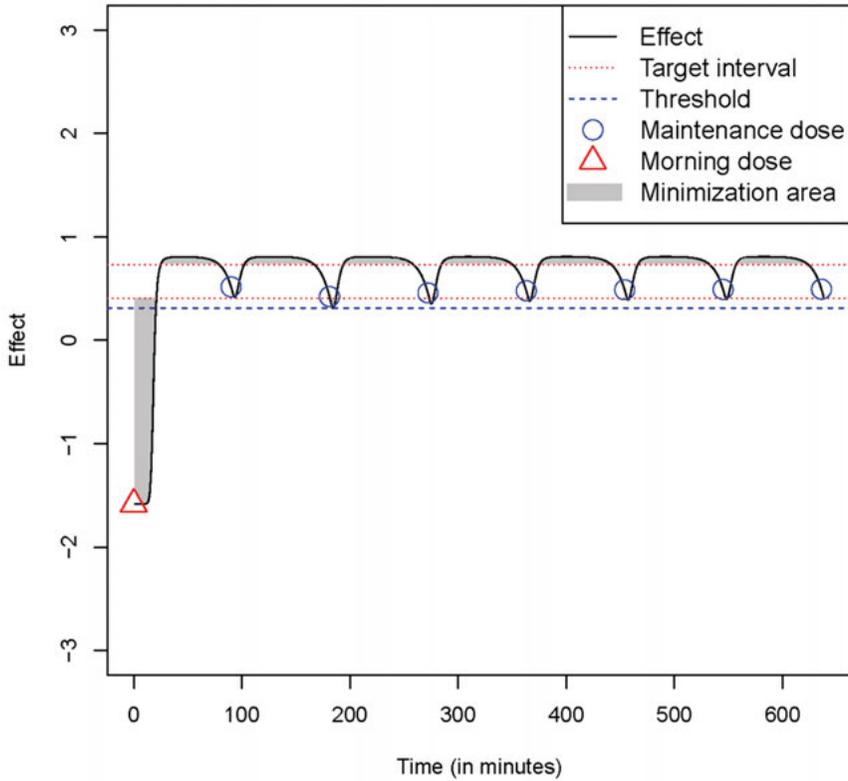


Figure 4. Dose optimization strategy of the algorithm for an example patient. The target range is indicated by the red (dotted) lines.

For input to the algorithm, individual models are used. The dosing algorithm produces an optimal dosing suggestion for a multitude of dosing intervals (time between two doses) for the individual models. Usually, a dosing interval is as low as 90 minutes, and as high as 240 minutes. The algorithm suggests one optimal dose for intervals within the 90-240-minute range, with a 10-minute interval increase. Thus, for 16 dosing intervals (90 min, 100 min, ..., 240 min), there would be 16 dosing suggestions, each one corresponding to a dosing interval.

For each dosing interval, in order to find the optimal dose, the algorithm performs repeated simulations in two steps. The first step is to find the maintenance dose, and the second step is to find the morning dose. For the maintenance dose, starting from a value of 1mg, the dose optimizer simulates the behavior of the dose-effect curve for 16 hours (960 minutes). The process is repeated for every dose until a maximum value (usually at 300 mg) with a one mg increase for the dose value. For each of the doses, the total area value

outside the target interval is stored (after the 600-minute mark, as the effect of a morning dose wears off at a maximum 10-hour period). That means that for a single interval, 300 different doses are evaluated and the dose that produces the minimum area outside the target interval is selected as the optimal. In the second step, the morning dose is evaluated in a similar way, but then the maximum dose is allowed to reach 400 mg, and the area evaluated is after the 20-minute mark (after receiving the dose the manifestation of effect has about a 20-minute delay, on average, so a choice before that could lead to overmedication of the patient). Finally, for every dosing interval, a combination of a morning dose and a maintenance dose is suggested as the optimal, the combination with the lowest sum of areas outside the target interval.

Data analysis

Paper I

The first step when analyzing raw signal data from sensors is usually feature extraction. For Dataset I, there were eight “raw” signals from the sensors, the three axes of acceleration, the three axes of rotation, and the overall magnitudes of rotation and acceleration (square root of the sum of squares of the three individual axes). From these signals 88 features were extracted, the mean, standard deviation and skewedness of each signal (24 features), approximate entropy (Pincus, 1991) features for each signal (eight features), and a discrete wavelet transform analysis which resulted in 48 features. Finally, eight features were calculated as the mean differences between the first half and the second half of the sensor readings for each signal. Even though feature extraction is a very important step in analysis of signals, it is not the focus of this thesis (or Paper I) and more details will not be provided.

The 88 extracted features were used to build predictive algorithms and evaluate whether patients motor statuses can be scored accurately through the pronation-supination test without clinical input. For Paper I, due the low ratio of observation/features, a principal components analysis (PCA) was performed to reduce the dimensionality of the dataset. After the PCA, 15 principal components were selected and used as input for four prediction algorithms: linear regression, decision trees, random forests, and a support vectors machine (SVM) model. The algorithms were trained in a supervised way, both for regression and classification. The objective for using the regression was to predict the subjects’ motor function score (clinical inputs were used as targets) in a continuous form on the TRS scale. The objective for using the classification was to test whether it is possible for the algorithms to recognize if the 20-second movements were performed by a subject or a healthy control. A bidirectional elimination method was used for feature selection for the SVM model. Pruning was used for the decision tree, and a backwards elimination method was used for the linear regression.

Since these algorithms can be prone to overfitting, the prediction performance of the algorithms was evaluated based on two cross-validation (CV) settings, a 10-fold setting and a leave-one-out (L1O) setting. In the first case the dataset was randomly split into 10 sets, and the algorithm was run ten times. For each run, nine sets were used for training and one set for testing and, after all runs, the results were averaged. For the L1O setting, the same

process was performed 19 times, each time leaving out the observations of a single subject during training, and then testing on that subject. The measures used to evaluate the performance were correlation values and root mean squared error values (RMSE). For the classification part, only 10-fold CV was tested and the measure of performance was classification accuracy.

Paper II

In Paper II, the individual model-fitting algorithm was applied on the chart-data (Dataset III) and for each subject included in the analysis an individual model was built. The Nelder-Mead algorithm was used for the optimization of the objective function, and for each model subject-specific parameters were calculated through this process. The models (the individual parameters calculated through the minimization process) were subsequently passed to the dose automation algorithm for continuous infusion of levodopa. Following minimization of the area (using the Nelder-Mead algorithm again) around the target value of 0, morning dose and infusion rate suggestions were produced for each subject. The dosing suggestions of the algorithm were compared to the subjects' discharge date dosing schedule (considered the gold standard). The measures used for the comparison were the relative error and the correlation values between the two dosing schedules.

Paper III

On this dataset the individual model-fitting method and the dosing algorithm for oral administration of levodopa were applied. First, the same 88 features as in Dataset I were calculated from the sensor readings of Dataset II, and PCA was applied. Next, the principal components were used as predictors in the SVM model built in Paper I to predict the subjects' status on the TRS without clinical input. Finally the sensor-based ratings from the series of pronation-supination tests that the subjects' performed were used to build individual models. The models were used as input to the dose automation algorithm for oral administration of levodopa, and for each subject a dosing schedule was produced. The supervising neurologist of this study produced dosing schedules for the subjects independently of the algorithmic suggestions, based on the PKG readings between the first and the second subject visit. The two dosing schedules were compared to each other (through t-test), and relative error and correlation values were the two measures used to evaluate the agreement.

Paper IV

For this analysis, the 88 features extracted from Dataset I were analyzed without PCA. The aim was to investigate if using a subset of the original features could produce similar, or better, results than the initial analysis. For this analysis the same four algorithms were used, but the predictors were original features instead of principal components. Two additional methods for dimension reduction/feature selection were tested, a step-wise regression method and Lasso regression. The results from those two methods were compared to the models in Paper I, based on the same measures (correlation, RMSE, classification accuracy).

Paper V

The main research question of Paper V was to check if patients show a wear-off in effect in the afternoon hours. For this analysis, the last day of admission from Dataset III was selected and analyzed. This is because stable plasma concentrations were most likely to have been achieved after dose adjustments (titration) up to the day of discharge, and any other medication the subjects received before admission was likely to be no longer effective. The TRS values between 09:00 hours and 12:00 hours were averaged, as well as TRS values between 12:00 and 17:00 hours. The mean values of the two periods were compared and the subjects were classified into three groups: meaningful increase for TRS value; meaningful decrease for TRS value; and no meaningful change for TRS. Furthermore, the mean TRS values for the morning and afternoon periods were analyzed to test if they were statistically significantly different. The same analysis was performed for the mean values of the infusion rates. Finally, linear mixed-effects models were fitted for each subject for the observations between 12:00 and 17:00 hours, in order to check if the linear models had a negative slope coefficient, which would provide more evidence of a wear-off in effect.

The same analyses were performed for Datasets IV and V, namely the mean values between morning hours and afternoon hours were compared, and linear mixed models were fitted in the observations during the afternoon hours. The subjects of these datasets were also classified into three subject groups, similar to Dataset III.

For each subject of dataset IV, the self-reported TRS values were averaged based on the time of the day an entry was made (after averaging there were only four observations for each subject: mean TRS at 08:00, 12:00, 16:00, and 20:00 hours). It is reasonable to assume that over the course of the three-year study period, the baseline TRS for the subjects changes as the disease progresses, but the overall trend of effect in the afternoon hours should remain unchanged. The TRS values at 08:00 and 12:00 were averaged as morning

value and the observations at 16:00 and 20:00 hours were averaged as afternoon value. Statistical tests were performed to check for significant differences between morning and afternoon hours. It is important to note that there were only four observations per day for this dataset and the mean values at 08:00 or 20:00 hours can influence the results considerably.

For the subjects in Dataset V, the observations at every hour were averaged for the three-day monitoring period (resulting in 10 observations per subject) and then the observations before 12:00 hours were averaged as morning value and the observations after 12:00 hours were averaged as afternoon value. In dataset V levodopa plasma concentrations values were available and the mean levodopa concentrations between the morning and the afternoon hours were also tested for significant differences.

Results

Paper I

When analyzing the data in Paper I, it was found that the calculated features can be good predictors when building data mining models for regression and classification. The prediction performance of the four examined models can be seen in Table 4, both for 10-fold CV, and for L1O CV. The SVM was selected as the model which showed a greater potential for generalization, but it should be noted that there is a large drop in performance between the 10-fold CV and L1O CV, indicating that making predictions on unseen data of unseen patients might be challenging. A visualization of the predictive performance is seen in Figure 5, where the results from the L1O CV are plotted against the mean TRS values of the three experts. The model predictions are annotated as TRIS (treatment response index from sensors)

Table 4 (© 2018 IEEE). Correlation and RMS error of the predictions of the machine learning algorithms from the 10-fold and the L1O CV.

	10-fold CV	10-fold CV	L1O CV	L1O CV
	RMSE	Correlation coefficient	RMSE	Correlation coefficient
SVM	0.73	0.82	1.14	0.49
Random forest	0.88	0.76	1.13	0.42
Decision tree	0.88	0.70	1.00	0.6
Linear regression	0.89	0.71	1.00	0.6

Furthermore, the model was tested for detecting levodopa sensitivity (through ANOVA tests) with results showing that, in general, the TRIS values showed a lower response than the TRS values. It was also tested how the TRIS values correlate with the UPDRS sub-items, and how those compare with TRS- UPDRS sub-items correlation. Overall, the correlations were not much different. One final test was carried out to evaluate the test-retest reliability of the TRIS between the un-medicated states of the subjects where it was found that TRIS

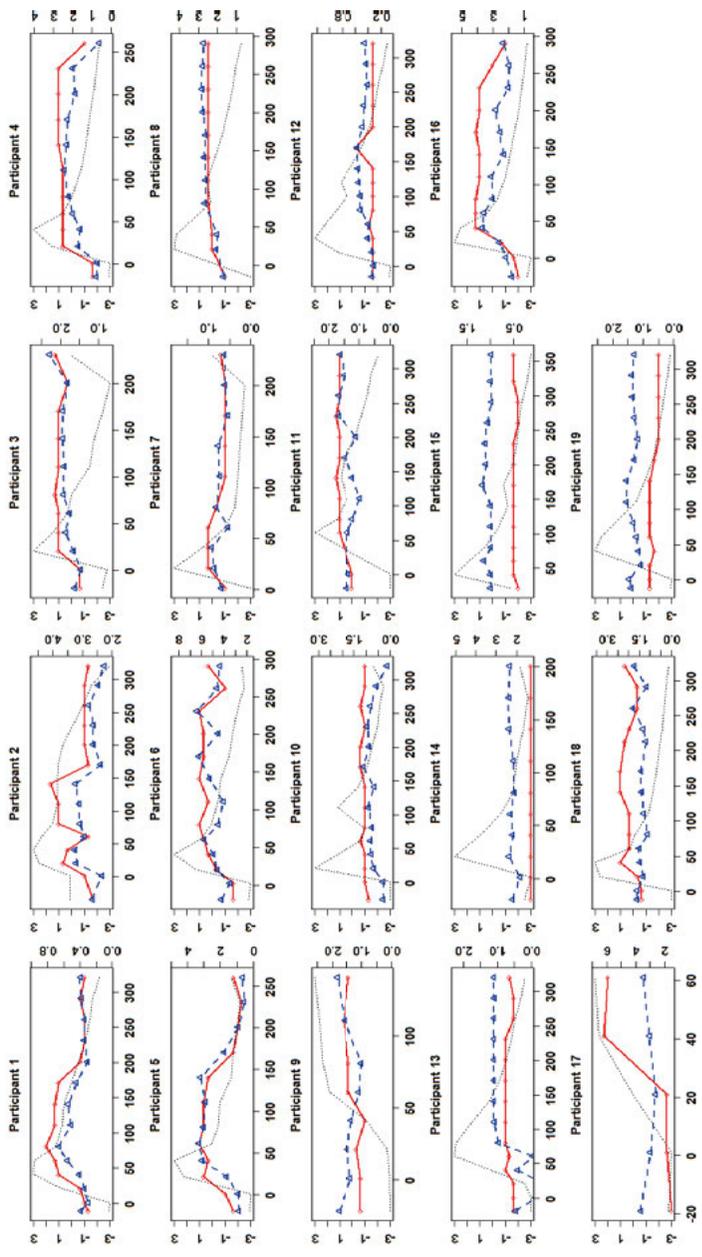


Figure 5 (© 2018 IEEE). The L10 predictions of the SVM model. The straight lines (circles) represent the TRS rating and the dashed lines (triangles) represent the TRIS values. These values are represented on the left axis, ranging from -3 to +3. The dotted lines show the levodopa blood plasma concentration (in mg/mL), and the values are shown on the right axis. The time (in minutes) from levodopa administration is shown on the x-axis.

predictions had ICC of 0.86. That value indicates that if a subject was to perform the test again with the same levodopa plasma concentration the algorithm would make consistent predictions. For the classification purposes, it was found that the SVM model trained for classification could correctly classify about 90% of the test occasions in a 10-fold CV setting.

Overall, the results showed that the use of an instrumental test to quantify the different states of the disease is possible, and the SVM model that was selected was used when analyzing the data in Paper III.

Paper II

The first finding of this study was the identification of the parameters that require individualization to fit individual models with a relatively small number of observations (8-10 per patient). Following sensitivity analysis, four parameters were selected for individualization from the mean PKPD model. These parameters are individualized for each patient, whereas the rest are fixed to population mean values.

The results from the automation algorithm for continuous infusion of levodopa are also presented in this paper. The algorithmic dosing suggestions and the optimized doses that the subjects had when leaving the clinic (Dataset III), were compared. It was found that the mean absolute relative errors between the infusion rates were low (10% with a 0.88 correlation value) and that the morning dose suggestions were consistently lower than the optimized morning doses from the clinic. This deviation was attributed to different optimization strategies between the algorithm and the treating personnel, since the algorithm is designed to avoid overmedication even in the morning hours, whereas in clinical practice the morning dose might be higher in order to reach an effect sooner. The individual models that were fitted can be seen in Figure 6, and a comparison between the morning doses and the infusion rates can be seen in Figure 7.

Paper III

This paper describes how the model individualization module described in Paper II can work with the sensor-based ratings developed in Paper I. The findings of the study were the evaluation of the ability to fit individual models, and how well the dosing suggestions, based on the individual models, compare to the dosing adjustments of the supervising neurologist.

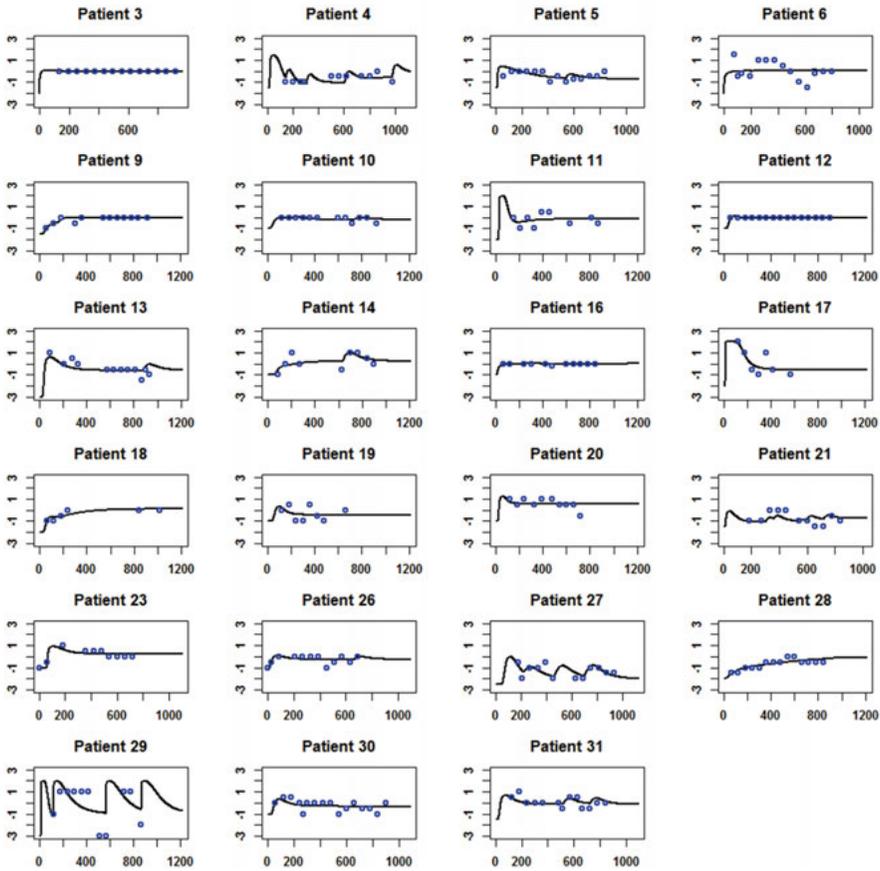


Figure 6. Individual fitted models. For each subject, one model was fitted with one day of information from the chart-data (doses, time of dosing, and ratings). On the y-axis, the response on the TRS is seen, and on the x-axis, the time (in minutes) is seen.

Out of 25 subjects who completed the second visit to the clinic, individual models were successfully fitted to 19 subjects. The only requirement for fitting a model in this setting was for the sequence of the subject-specific ratings during the levodopa challenge test to increase from a baseline value to a peak value, and then return to the baseline value at the end of the levodopa challenge. The reason six individual models could not be fitted is because those sequences of ratings did not show this behavior. These results compared favorably with the sequences of TRS ratings regarding the ability to fit individual models (11 subjects), mainly because with the TRS consensus ratings there was no return to baseline at the end of the cycle (the effect plateaued after reaching peak effect).

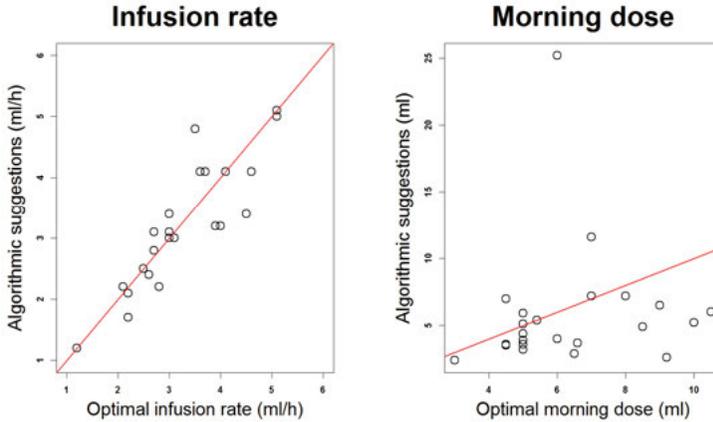


Figure 7. Left panel: infusion rate comparison. Right panel: morning dose comparison. In both panels, optimal dose vs. algorithmic dose is plotted with a line added going through the origin.

In order to evaluate the performance of the individual models (and the methodology in general), the dosing schedules of the automation algorithm were compared with the PKG-based dosing adjustments (Johansson et al., 2018), and, as shown in Figure 8, there was good agreement between the two. Furthermore, the t-tests showed no significant statistical differences, and the relative errors between the doses were 21% and 12.5% for morning and maintenance dose, respectively, with a correlation of 0.8 and 0.95.

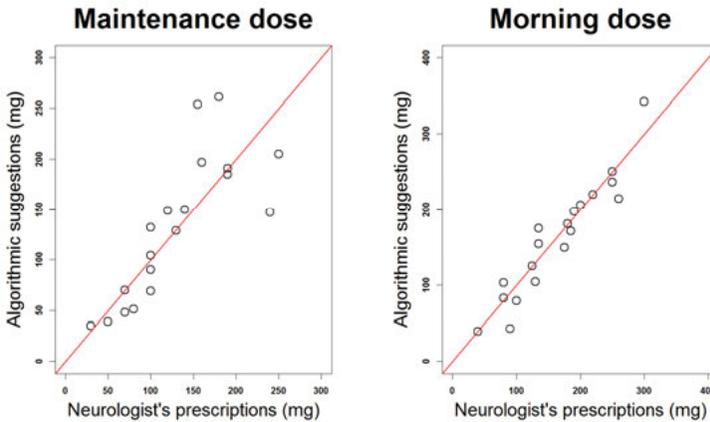


Figure 8. Left panel: maintenance dose comparison. Right panel: morning dose comparison. In both panels, optimal dose vs. algorithmic dose is plotted with a line added going through the origin.

Paper IV

Given that in Paper I the SVM model predictions had clear limitations when assessing values at the bounds of the TRS scale, Dataset I was analyzed again to try and address the limitations. Two additional methods were used in this analysis; a step-wise regression method, and Lasso regression. The methods were evaluated on both the 10-fold CV and the L1O setting. It was found that the SVM model with original features selected through the stepwise regression approach produced TRIS ratings that had better agreement with the TRS ratings, and the improvement was marginal in the 10-fold CV setting (0.82 correlation increased to 0.84), greater on the L1O setting (0.49 correlation increased to 0.65), but the performance on the extremes did not improve (following visual inspection). However, the new TRIS values had higher sensitivity to levodopa, compared to the old TRIS values, as seen in Figure 9. The results from the analysis can be seen in Table 5.

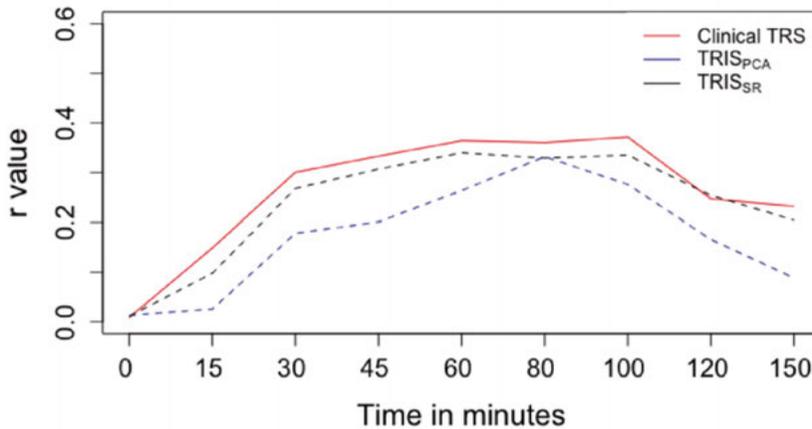


Figure 9 (© 2018 IEEE). Sensitivity to treatment of TRS, TRIS_{PCA} and TRIS_{SR}. The r value on the y-axis are the effect sizes as acquired from ANOVA tests between the baseline observation and the subsequent observations.

Table 5 (© 2018 IEEE). Correlation and root mean squared error of the machine learning methods from the 10-fold and the L1O CV when using PCA, step-wise regression (SR), and Lasso

	10-fold CV	10-fold CV	L1O CV	L1O CV
	RMSE	Correlation coefficient	RMSE	Correlation coefficient
SVM-PCA	0.73	0.82	1.14	0.49
SVM-SR	0.69	0.84	0.98	0.65
SVM-Lasso	0.95	0.68	1.07	0.41
RF-PCA	0.88	0.76	1.13	0.42
RF-SR	0.72	0.83	1.00	0.58
RF-Lasso	0.77	0.82	1.00	0.60
DT-PCA	0.88	0.70	1.00	0.60
DT-SR	1.01	0.60	1.08	0.45
DT-Lasso	0.99	0.61	1.09	0.44
LR-PCA	0.89	0.71	1.00	0.60
LR-SR	0.81	0.77	0.98	0.65
LR-Lasso	0.81	0.78	1.09	0.52

Paper V

The three datasets analyzed in this paper had ratings that were from three different sources: Dataset II had TRS ratings from nurses on the ward; Dataset IV has subjects' self-reported TRS ratings; and Dataset IV had TRS ratings, as scored by a movement disorders expert. During the analysis of this paper an interesting pattern was discovered regarding the effect of continuous levodopa infusion during the afternoon hours. In all three datasets there was a drop in effect in the afternoon hours for about half of the subjects, and the results can be seen in Table 6. The proportion of subjects with a large decrease in TRS ranges from 43% in Dataset II, to 58% in Dataset V.

For all three datasets, it was detected that there were statistically significant differences between morning TRS values and afternoon TRS values. What should be noted here is that in Dataset II, the amounts of levodopa the subjects received during the morning and afternoon hours were not significantly different, and in Dataset IV, the levodopa plasma concentrations during the morning and afternoon hours were not significantly different. That would indicate that even with the same levodopa plasma concentrations, the patients exhibit different effect values. The PKPD models currently used do not accommodate for such changes, given that they are either deterministic, or they do not have a time coefficient.

Table 6. *Classification of subjects based on values of increase and decrease of the TRS during the afternoon hours.*

	n (%) of subjects without any meaningful change in TRS during the latter part of the day	n (%) of subjects with a decrease of at least 0.2 TRS points during the latter part of the day	n (%) of subjects with an increase of at least 0.2 TRS points during the latter part of the day
Dataset III (N=28)	13 (46 %)	12 (43 %)	3 (10 %)
Dataset IV (N=58)	18 (31 %)	29 (50 %)	11 (19 %)
Dataset V (N=12)	3 (25 %)	7 (58 %)	2 (17 %)
Total (N=98)	34 (35 %)	48 (49 %)	16 (16 %)

Conclusions

This thesis presented a comprehensive solution that describes how one can proceed from data collection, data processing, and data analysis to data-driven decision support for Parkinson's disease treatment individualization. A few steps were included in this process and the highlights of the work were shown to provide a holistic idea of how this method works, and under which conditions.

The first step in this study was to obtain data that contained the necessary information for treatment individualization. Two ways were explored: one by means of medical chart-data, and the other by wearable sensors. For the future application of this approach, sensor data are preferable since they do not require the presence of clinical experts, or the patients' admission to hospitals. The clinical data contain ratings of the patients' motor function in response to the dose received, as opposed to the sensors which only capture raw signals. Thus, in Paper I a method is described to show how raw signals can be converted to objective ratings. The results of paper I indicated that it is possible to monitor the disease motor states in a holistic way (research Question I), but the method also had limitations, meaning that it needs further refinement before being used in clinical practice.

The second step in the study was to explore how individual patient ratings could be used to model patient-specific dose-response behaviors. This was done by developing a method that allows the construction of patient-specific profiles (or individual models) with only a few ratings, as was the case in the datasets. The method presented in Paper II was robust and allowed for the individualization of a population PKPD model for levodopa intake (research Question 2). It should be noted that more solutions to create patient-specific profiles could be possible.

The third step in the process was to develop the dose automation algorithms presented in the methodological framework section. Both algorithms were evaluated, the first with chart-data, and the second with data from wearables. The results from the two algorithms demonstrated the feasibility and the potential of the solution. Both algorithms used the individual patient profiles for the dosing automation, demonstrating how the profiles can be used in practice (research Question 3).

Finally, in relation to treatment individualization, it was debated whether the PKPD models currently used are appropriate. There were reports in the literature that the patients' effect during the afternoon hours tends to be lower,

even with stable levodopa plasma concentrations. This question was investigated in Paper V, and it was found that for about half of the subjects in the study that statement was true (research Question 4).

In this thesis, emphasis was given to data collection, data description, and data processing. There were two main reasons for that choice. First, it is necessary to describe all the steps of the process to make the studies as reproducible as possible. I hope these descriptions facilitated the reader's understanding of the challenges to be overcome and the decisions made during data processing. Second, this thesis is in the field of Micro-data analysis, where understanding data is a large part of the contribution. Comprehending the problem is an extensive part of proposing a solution, but appreciating the resources and adapting them accordingly is as an important part of the solution as any.

Discussion

In a recent review paper (Odin et al., 2018), a panel of 11 internationally recognized movement disorder experts met over a two-year period (2014-2016) and discussed four principles for objective measurements that were answered in a series of seven questions and answers. For most of the questions there was panel consensus and clear agreement on what should guide the use of sensors in monitoring PD. Related to the questions laid out in the introduction of this thesis, they agreed that an ideal device should continuously “measure complications and therapy responsive features while PwP undertake their usual daily activities”. That opinion would address the issue of long-term vs. short-term monitoring, indicating a leaning towards a “long-term” monitoring period. In this present thesis it was found that short-term monitoring with instrumental tests could also be beneficial and contain information. I would agree that the more data collected the more information one is likely to extract, but I am not convinced that passive monitoring is the only solution. I do, however, understand the value of long-term monitoring, and as a future course, research could be directed into examining how well the dosing algorithms will work with passive monitoring devices.

There is an abundance of published research where data are collected during passive monitoring or during instrumental tests, or by a combination of both. My personal view is that asking the patients to wear a device continuously, and at the same time ask them to do instrumental tests is not practicable, since it means the patient’s life is made to revolve around data collection. Even though the position of devices is not discussed in Odin et al. (2018), in other published research multiple sensor positions are examined. It is worth noting that the two FDA-approved devices, PKG only uses a wrist worn sensor, and Kinesia uses wrist and ankle bands. My view is aligned with this approach, since, on the one hand, in Paper I we found that collecting data through wrist worn sensors might be enough, and, on the other hand, the use on multiple sensors is extra work for the patients. This view is further supported by the results from Thomas et al. (2017), where data from all four sensors from Dataset I were analyzed, and it was found that the wrist sensors had the ability to better analyze the patients’ motor status, and adding foot sensors did not meaningfully improve the performance of the prediction models. I feel it is unrealistic to ask patients to wear head and waist sensors, or even suits, in order to collect data. This view is also reflected in publications examining the patients’ views on wearables, where sometimes patients also object to long-

term monitoring, which leaves instrumental tests as the only viable alternative. The views of patients on wearable sensors and devices have not been extensively reviewed in the literature; however, some studies have shown negative disposition against the continuous monitoring (or continuous long-term monitoring). An example of this is in the research of Ozanne et al. (2018), where a focus group study is presented and the results are discussed.

Even though this thesis study did not investigate sensor position or length of monitoring, I feel this discussion is well placed, when considering the potential utilization of the methods developed. Their application will be in conjunction with objective rating systems, and I think it is necessary to point out that the methods developed should be further adapted to accommodate different monitoring systems. I also believe that the results of the thesis make a contribution to the wearables debate, since instrumental tests during levodopa challenge tests produced promising results.

Instrumental monitoring should, and usually is performed during a levodopa challenge test, as in Paper I, since a full levodopa cycle should be monitored to extract relevant information. This is even more important when considering the purpose of the experiment, which is usually for dosing adjustments, whether algorithmic, as in Papers II and III, or expert-based, as in clinical practice. This claim is further supported by the results of Paper II, where I believe that had the data been collected during a levodopa cycle, the agreement between the morning doses would be higher (Figure 7, right panel).

This thesis can be described as a proof-of-concept work that shows how the methods could be used with further development. The proposed solution was robust, and as a proof-of-concept study it was successful. Overall, if one were to place the contribution of the study in any field, the most appropriate place would be to view it as a contribution to the ongoing discussion regarding the use of wearable sensors in Parkinson's disease and their potential application. Even though objective ratings were not used in Paper II, one can envision how the method would be adapted without many modifications. The thesis focus is not wearable technology, but rather how wearables could be utilized for dose automation.

Ethical considerations

For this thesis, collection and handling of personal data was necessary. When collecting personal data (from patients or otherwise) it is imperative, firstly, to obtain consent, and, secondly, to protect the data collected. All subjects participating in the studies in this thesis had given written consent for their data to be collected and used for research purposes. Furthermore, ethical permission was obtained from the appropriate local committees for all the studies (with the exception of Dataset IV, where the Regional Ethical Committee determined that ethical approval was not required).

When handling and processing data, only signal readings and subject characteristics, without personal identifiers, were available in Dataset I. For Dataset II, the scanned copies from the clinic did not contain any identifiers of subjects. For Dataset III, data was sent through an online secure server from the study site. For Datasets IV and V, no identifiers were available.

This research not only considers personal data, but could potentially have an important impact on the medication schedules of patients, and therefore on the quality of their lives. That realization was important to keep in mind at every step of the process, while understanding that behind the data there are individuals who suffer from a disease. The wellbeing of each patient is the most important consideration, thus, in future studies the algorithm suggestions given here should be carefully examined by movement disorder experts before being prescribed to a patient. It is hoped that in future iterations the algorithmic dosing suggestions will not require clinical approval, but even then, there should be specific dosing thresholds to protect patients, when dosing suggestions seem unrealistic.

Limitations

The method presented in Paper I for rating the patient's status was accurate but had limitations, particularly when scoring values close to the bounds of the TRS. A few methods were explored to address those limitations, but, overall, the performance of the algorithm on the extremes did not improve. Paper IV describes a different feature selection method for the algorithm, but there was no obvious improvement of the TRIS performance. Furthermore, resampling methods with the same dataset were explored but the predictive performance on the extremes did not improve. These results were presented as a Master's Thesis in Business Intelligence which I supervised (Dutra, 2018). Even though the algorithm had limitations, its performance was validated in a second clinical study (the data used in Paper III), and, overall, the TRIS from that study showed reduced, but similar, RMSE and correlation to the TRS (Johansson et al., 2019). Having lower performance on a validation study is common, and possibly can be expected when applying the same method to an independent population. There were also different parameters between the two populations, namely, the amount of levodopa patients received (150% compared to 120%), which changed the distribution of the TRS between the two studies, as seen in Figure 1. Another major difference was that in Dataset I there were three movement disorder experts who rated independently, while in the Dataset III, there were two movement disorder experts who reached a consensus rating. This changed the distribution from somewhat continuous to an ordinal scale and, since the SVM model was trained for regression, the results from Johansson et al. (2019) were further influenced.

One major limitation of this thesis is that the dosing algorithm for oral administration of levodopa does not currently offer one optimal dosing schedule, but, instead, gives a suggestion of dosing schedules based on the number of dosing intervals investigated. More research needs to be carried out in this direction. One method was investigated comparing the areas outside the target interval between the different dosing schedules, but this was discarded since the results were not very informative and their interpretation was challenging. This research needs to be undertaken from scratch, and having the data used in Paper III could be an appropriate start for this research. Another study would be required, though, to validate that the dosing schedules suggested are appropriate.

In Paper V, even though there were indications of a drop in effect during the afternoon hours, no explanation was presented to address this phenomenon, other than speculation of possible reasons. Further study is necessary to investigate why this occurs, which should be part of future studies examining the system dynamics of levodopa.

Future research

The main goal of future research for this project could be to validate the performance of the dosing algorithms, in terms of ability to provide dosing schedules that improve the patient's motor functions, and to explore if continuous objective monitoring can also be used for making dosing adjustments. Adapting the dosing algorithm to passive measurements is important, as the system could then be made available in the home environment.

As a future direction, unsupervised machine-learning algorithms could also be explored to analyze the data from passive patient monitoring. The simulation-based dosing algorithm could be further adapted, and data mining techniques could be used to recognize dosing patterns and mimic expert dosing decisions. Future work could focus on evaluation, refinement, and validation of the system.

Moreover, the system currently only works with TRS, a scale that is not widely used outside Sweden. More research is necessary to adapt the system and create a methodology that creates a patient-specific model using ratings from other scales, such as the UPDRS.

Finally, a new PKPD model that allows for a time-of-day covariate could be explored to model the drop off effect during afternoon hours.

Acknowledgments

This thesis was carried out at Dalarna University with funds from The Swedish Knowledge Foundation (KK-stiftelsen) and VINNOVA. It would not have been possible without the contribution and support of all the people who helped me through this process.

First of all, I would like to thank my supervisors for their advice. My primary supervisor, Associate Professor Jerker Westin, for your great patience and the ability to explain your vision for the path you laid out for me. You gave me room to struggle, succeed, question my English language skills, and ultimately grow into a better scientist and person. My secondary supervisor Associate Professor Moudud Alam for your intelligence and your expertise in all matters statistics. You showed me how much I don't know and the long way I have to cover to grow into an independent researcher.

Thank you to the two professors, Kenneth and Mark. Kenneth, your work ethic is inspiring, and our trip to Baltimore was memorable. Mark, thank you for keeping me from spending all my money in Las Vegas.

To all my colleagues in the statistics corridor, Lars, Mengjie, Xiaoyun, Yujiao, Ross, Reuben, and Ola, thank you for making our subject work smoothly.

Thank you to all my colleagues at Dalarna University who made the work environment cozy and relaxed. Special thanks to Catia, for being so honest and friendly, and Pascal, for also being so honest and friendly. Hasan, thank you for the good discussions.

Thank you to all my co-authors. Filip, your British wit is uncanny, even though you are not British. I truly enjoyed working with you. Dag, I feel you acted like a ghost supervisor, and your support is greatly appreciated. Mevludin, you have been a pillar of consistency and you never lose focus. Dongni, working and travelling with you has been one of the highlights of my PhD studies. Marina, thank you for being so well-versed and helpful when required. Special thanks also to Nicola and Alessio. My time with you inspired me to work towards a PhD.

To my fellow PhD students, Sami, Charlie, Vijay, and Asif. Murshid, thank you for sharing an office with me and being the only person in Sweden who watches Doctor Who. Åsa, thank you for being such a good friend and also so supportive and helpful.

Thank you to my childhood friends, Achilleas, Christos, Giorgos, Konstantinos, Tasos, and Vasilis. My doing a PhD had nothing to do with them, but at this point I am playing with house money, so I might as well name them.

Claudia and Andres, you made our time in Borlänge fun and enjoyable. Mihaela, thank you for being a great human travel companion and an amazing friend. Aysuhan, you live halfway around the world but you are in my heart. Gülnur, you are the best friend I could ask for.

Eleni, you are my moral compass. Thank you for tolerating me. I hope our journey together is only just beginning.

Finally, I would like to thank my parents for always doing their best, believing in me, and their unconditional love.

Περίληψη – Abstract (in Greek)

Η νόσος του Πάρκινσον είναι η δεύτερη πιο συχνή νευροεκφυλιστική ασθένεια. Η λεβοντόπα χρησιμοποιείται κυρίως για τη διαχείριση των κινητικών συμπτωμάτων της νόσου. Ωστόσο, η μακροχρόνια χρήση λεβοντόπα και η πρόοδος της ασθένειας προκαλεί μειωμένη αποτελεσματικότητα και παρενέργειες. Όταν συμβαίνει αυτό, απαιτούνται ακριβή εξατομικευμένα προγράμματα δοσολογίας.

Αυτή η διδακτορική διατριβή στον τομέα της ανάλυσης μικροδεδομένων προτείνει μια πλήρη λύση για την αυτοματοποίηση της φαρμακολογικής διαχείρισης της νόσου του Πάρκινσον, με λεβοντόπα, και προσφέρει μια εκτίμηση για τη φαρμακοδυναμική λειτουργία της. Για το σκοπό αυτό, προτείνεται ένας αλγόριθμος που παράγει αντικειμενικές εκτιμήσεις της κινητικής λειτουργίας των ασθενών, μέσω φορητών αισθητήρων, αναπτύσσεται μια μέθοδος για την κατασκευή εξατομικευμένων προφίλ για τους ασθενείς, και παρουσιάζονται δύο αλγόριθμοι δοσολογίας, ένας για την στοματική, και ένας για την εντερική χορήγηση λεβοντόπα. Για την ανάπτυξη των μεθόδων και την αξιολόγηση της απόδοσης των προτεινόμενων αλγορίθμων χρησιμοποιήθηκαν δεδομένα από πέντε διαφορετικές πηγές.

Οι αλγόριθμοι αυτοματοποίησης της δοσολογίας μπορούν να λειτουργήσουν τόσο με κλινικές όσο και με αντικειμενικές εκτιμήσεις (μέσω φορητών συσκευών), και η εφαρμογή τους αξιολογήθηκε σε συνάρτηση με προσαρμογές δοσολογίας από εμπειρογνώμονες σε κινητικές διαταραχές. Οι υποδείξεις χορήγησής των δύο αλγορίθμων ήταν παρόμοιες με αυτές των εμπειρογνομόνων, επιδεικνύοντας τις δυνατότητες της προτεινόμενης μεθόδου.

Ο αλγόριθμος αντικειμενικών εκτιμήσεων είχε καλή αξιοπιστία εξέτασης-επανεξέτασης και η εφαρμογή του κατά τη διάρκεια μιας κλινικής μελέτης ήταν επιτυχής. Επιπλέον, η μέθοδος της εξατομικευμένων προφίλ ασθενών ήταν ευπροσάρμοστη και δούλεψε καλά με τον αλγόριθμο αντικειμενικών εκτιμήσεων. Τέλος, διεξήχθη μια μελέτη που έδειξε ότι περίπου οι μισοί ασθενείς που υποβάλλονται σε θεραπεία με λεβοντόπα παρουσιάζουν μειωμένη ανταπόκριση κατά τις απογευματινές ώρες, επισημαίνοντας την ανάγκη για πιο ακριβή μοντελοποίηση της φαρμακοδυναμικής της λεβοντόπα.

References

- Ahlskog, J. E., & Muenter, M. D. (2001). Frequency of levodopa-related dyskinesias and motor fluctuations as estimated from the cumulative literature. *Movement Disorders, 16*(3), 448-458.
- Caramia, C., Torricelli, D., Schmid, M., Muñoz-Gonzalez, A., Gonzalez-Vargas, J., Grandas, F., & Pons, J. L. (2018). IMU-Based Classification of Parkinson's Disease From Gait: A Sensitivity Analysis on Sensor Location and Feature Selection. *IEEE Journal of Biomedical and Health Informatics, 22*(6), 1765-1774.
- Connolly, B. S., & Lang, A. E. (2014). Pharmacological treatment of Parkinson disease: a review. *Jama, 311*(16), 1670-1683.
- Dorsey, E., Constantinescu, R., Thompson, J. P., Biglan, K. M., Holloway, R. G., Kieburtz, K., ... & Tanner, C. M. (2007). Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. *Neurology, 68*(5), 384-386.
- Dorsey, E., Vlaanderen, F. P., Engelen, L. J., Kieburtz, K., Zhu, W., Biglan, K. M., ... & Bloem, B. R. (2016). Moving Parkinson care to the home. *Movement Disorders, 31*(9), 1258-1262.
- Dorsey, E. R., George, B. P., Leff, B., & Willis, A. W. (2013). The coming crisis: Obtaining care for the growing burden of neurodegenerative conditions. *Neurology, 10*-1212.
- Dutra, F. (2018) Evaluation of Calibration Methods to Adjust for Infrequent Values in Data for Machine Learning.
- Ebersbach, G., Baas, H., Csoti, I., Müngersdorf, M., & Deuschl, G. (2006). Scales in Parkinson's disease. *Journal of neurology, 253*, iv32-iv35.
- Espay, A. J., Bonato, P., Nahab, F. B., Maetzler, W., Dean, J. M., Klucken, J., ... & Reilmann, R. (2016). Technology in Parkinson's disease: Challenges and opportunities. *Movement Disorders, 31*(9), 1272-1282.
- Fabbrini, G., Brotchie, J. M., Grandas, F., Nomoto, M., & Goetz, C. G. (2007). Levodopa-induced dyskinesias. *Movement Disorders, 22*(10), 1379-1389.
- Fenoy, A. J., & Simpson Jr, R. K. (2014). Risks of common complications in deep brain stimulation surgery: management and avoidance. *Journal of Neurosurgery, 120*(1), 132-139.
- Griffiths, R. I., Kotschet, K., Arfon, S., Xu, Z. M., Johnson, W., Drago, J., ... & Horne, M. K. (2012). Automated assessment of bradykinesia and dyskinesia in Parkinson's disease. *Journal of Parkinson's Disease, 2*(1), 47-55.
- Giuffrida, J. P., Riley, D. E., Maddux, B. N., & Heldman, D. A. (2009). Clinically deployable Kinesia™ technology for automated tremor assessment. *Movement Disorders, 24*(5), 723-730.
- Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., ... & Dubois, B. (2008). Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Movement Disorders, 23*(15), 2129-2170.

- Heldman, D. A., Giuffrida, J. P., Chen, R., Payne, M., Mazzella, F., Duker, A. P., ... & Espay, A. J. (2011). The modified bradykinesia rating scale for Parkinson's disease: reliability and comparison with kinematic measures. *Movement Disorders*, 26(10), 1859-1863.
- Jankovic, J. (2005). Motor fluctuations and dyskinesias in Parkinson's disease: clinical manifestations. *Movement Disorders*, 20(S11).
- Jankovic, J. (2008). Parkinson's disease: clinical features and diagnosis. *Journal of Neurology, Neurosurgery & Psychiatry*, 79(4), 368-376.
- Jayachandran, D., Rundell, A. E., Hannemann, R. E., Vik, T. A., & Ramkrishna, D. (2014). Optimal chemotherapy for leukemia: a model-based strategy for individualized treatment. *PLoS one*, 9(10), e109623.
- Johansson, D., Thomas, I., Ericsson, A., Johansson, A., Medvedev, A., Memedi, M., Nyholm, D., Ohlsson, F., ... Westin, J., & Bergquist, F. (2019). Evaluation of a sensor algorithm for motor state rating in Parkinson's disease. *Draft*.
- Johansson, D., Ericsson, A., Johansson, A., Medvedev, A., Nyholm, D., Ohlsson, F., ... & Bergquist, F. (2018). Individualization of levodopa treatment using a microtablet dispenser and ambulatory accelerometry. *CNS Neuroscience & Therapeutics*, 24(5), 439-447.
- Johansson, D., Malmgren, K., & Murphy, M. A. (2018). Wearable sensors for clinical applications in epilepsy, Parkinson's disease, and stroke: a mixed-methods systematic review. *Journal of neurology*, 1-13.
- Kubota, K. J., Chen, J. A., & Little, M. A. (2016). Machine learning for large-scale wearable sensor data in Parkinson's disease: Concepts, promises, pitfalls, and futures. *Movement Disorders*, 31(9), 1314-1326.
- Leelarathna, L., Dellweg, S., Mader, J. K., Allen, J. M., Benesch, C., Doll, W., ... & Michalewski, L. (2014). Day and night home closed-loop insulin delivery in adults with type 1 diabetes: three-center randomized crossover study. *Diabetes Care*, 37(7), 1931-1937.
- Lopane, G., Mellone, S., Chiari, L., Cortelli, P., Calandra-Buonaura, G., & Contin, M. (2015). Dyskinesia detection and monitoring by a single sensor in patients with Parkinson's disease. *Movement Disorders*, 30(9), 1267-1271.
- Maetzler, W., Klucken, J., & Horne, M. (2016). A clinical view on the development of technology-based tools in managing Parkinson's disease. *Movement Disorders*, 31(9), 1263-1271.
- Mancini, F., Comi, C., Oggioni, G. D., Pacchetti, C., Calandrella, D., Moja, M. C., ... & Lacerenza, M. (2014). Prevalence and features of peripheral neuropathy in Parkinson's disease patients under different therapeutic regimens. *Parkinsonism & Related Disorders*, 20(1), 27-31.
- Mancini, M., Salarian, A., Carlson-Kuhta, P., Zampieri, C., King, L., Chiari, L., & Horak, F. B. (2012). ISway: a sensitive, valid and reliable measure of postural control. *Journal of Neuroengineering and Rehabilitation*, 9(1), 59.
- Memedi, M., Nyholm, D., Johansson, A., Pålhagen, S., Willows, T., Widner, H., ... & Westin, J. (2015). Validity and responsiveness of at-home touch screen assessments in advanced Parkinson's disease. *IEEE Journal of Biomedical and Health Informatics*, 19(6), 1829-1834.
- Merello, M., Hughes, A., Colosimo, C., Hoffman, M., Starkstein, S., & Leiguarda, R. (1997). Sleep benefit in Parkinson's disease. *Movement Disorders*, 12(4), 506-508.
- Jane, Y. N., Nehemiah, H. K., & Arputharaj, K. (2016). A Q-backpropagated time delay neural network for diagnosing severity of gait disturbances in Parkinson's disease. *Journal of Biomedical Informatics*, 60, 169-176.

- Nyholm, D., Askmark, H., Gomes–Trolin, C., Knutson, T., Lennernäs, H., Nyström, C., & Aquilonius, S. M. (2003). Optimizing levodopa pharmacokinetics: intestinal infusion versus oral sustained-release tablets. *Clinical Neuropharmacology*, 26(3), 156-163.
- Nyholm, D., Remahl, A. N., Dizdar, N., Constantinescu, R., Holmberg, B., Jansson, R., ... & Askmark, H. (2005). Duodenal levodopa infusion monotherapy vs oral polypharmacy in advanced Parkinson disease. *Neurology*, 64(2), 216-223.
- Nyholm, D., and Stepien, V. (2014). Levodopa fractionation in Parkinson's disease. *Journal of Parkinson's disease*, 4(1), 89-96.
- Odin, P., Chaudhuri, K. R., Volkmann, J., Antonini, A., Storch, A., Dietrichs, E., ... & Bergquist, F. (2018). Viewpoint and practical recommendations from a movement disorder specialist panel on objective measurement in the clinical management of Parkinson's disease. *NPJ Parkinson's disease*, 4(1), 14.
- Olanow, C. W., Obeso, J. A., & Stocchi, F. (2006). Continuous dopamine-receptor treatment of Parkinson's disease: scientific rationale and clinical implications. *The Lancet Neurology*, 5(8), 677-687.
- Othman, A. A., & Dutta, S. (2014). Population pharmacokinetics of levodopa in subjects with advanced Parkinson's disease: levodopa-carbidopa intestinal gel infusion vs. oral tablets. *British Journal of Clinical Pharmacology*, 78(1), 94-105.
- Ozanne, A., Johansson, D., Hällgren Graneheim, U., Malmgren, K., Bergquist, F., & Alt Murphy, M. (2018). Wearables in epilepsy and Parkinson's disease—A focus group study. *Acta Neurologica Scandinavica*, 137(2), 188-194.
- Papapetropoulos, S. S. (2012). Patient diaries as a clinical endpoint in Parkinson's disease clinical trials. *CNS Neuroscience & Therapeutics*, 18(5), 380-387.
- Parkinson, J. (2002). An essay on the shaking palsy. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 14(2), 223-236.
- Patel, S., Lorincz, K., Hughes, R., Huggins, N., Growdon, J., Standaert, D., ... & Bonato, P. (2009). Monitoring motor fluctuations in patients with Parkinson's disease using wearable sensors. *IEEE Transactions on Information Technology in Biomedicine*, 13(6), 864-873.
- Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences*, 88(6), 2297-2301.
- Ramaker, C., Marinus, J., Stiggelbout, A. M., & Van Hilten, B. J. (2002). Systematic evaluation of rating scales for impairment and disability in Parkinson's disease. *Movement Disorders*, 17(5), 867-876.
- Simon, N., Viallet, F., Boulamery, A., Eusebio, A., Gayraud, D., & Azulay, J. P. (2016). A combined pharmacokinetic/pharmacodynamic model of levodopa motor response and dyskinesia in Parkinson's disease patients. *European Journal of Clinical Pharmacology*, 72(4), 423-430.
- Senek, M., Aquilonius, S. M., Askmark, H., Bergquist, F., Constantinescu, R., Ericsson, A., ... & Spira, J. (2017). Levodopa/carbidopa microtablets in Parkinson's disease: a study of pharmacokinetics and blinded motor assessment. *European Journal of Clinical Pharmacology*, 73(5), 563-571.
- Sensidose AB, Authors: Ilias Thomas, Mevluding Memedi, and Jerker Westin (UK), "SYSTEMS FOR EVALUATING DOSAGE PARAMETERS", GB1708624.0, 2017
- Stamate, C., Magoulas, G. D., Küppers, S., Nomikou, E., Daskalopoulos, I., Jha, A., ... & Iannone, M. (2018). The cloudUPDRS app: A medical device for the clinical assessment of Parkinson's Disease. *Pervasive and mobile computing*, 43, 146-166.

- Struys, M. M., De Smet, T., Glen, J. I. B., Vereecke, H. E., Absalom, A. R., & Schnider, T. W. (2016). The history of target-controlled infusion. *Anesthesia & Analgesia*, *122*(1), 56-69.
- Timpka, J., Henriksen, T., & Odin, P. (2016). Non-oral Continuous Drug Delivery Techniques in Parkinson's Disease: For Whom, When, and How?. *Movement Disorders Clinical Practice*, *3*(3), 221-229.
- Tsipouras, M. G., Tzallas, A. T., Rigas, G., Tsouli, S., Fotiadis, D. I., & Konitsiotis, S. (2012). An automated methodology for levodopa-induced dyskinesia: assessment based on gyroscope and accelerometer signals. *Artificial Intelligence in Medicine*, *55*(2), 127-135.
- Tomlinson, C. L., Stowe, R., Patel, S., Rick, C., Gray, R., & Clarke, C. E. (2010). Systematic review of levodopa dose equivalency reporting in Parkinson's disease. *Movement Disorders*, *25*(15), 2649-2653.
- Thomas, I. (2017). Optimizing levodopa dosing routines for Parkinson's disease (Licentiate dissertation, Högskolan Dalarna).
- Thomas, I., Bergquist, F., Constantinescu, R., Nyholm, D., Senek, M., & Memedi, M. (2017). Using measurements from wearable sensors for automatic scoring of Parkinson's disease motor states: Results from 7 patients. In *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*.
- Uem, J. M., Maier, K. S., Hucker, S., Scheck, O., Hobert, M. A., Santos, A. T., ... & Maetzler, W. (2016). Twelve-week sensor assessment in Parkinson's disease: Impact on quality of life. *Movement Disorders*, *31*(9), 1337-1338.
- Valldeoriola, F., Puig-Junoy, J., Puig-Peiró, R., & Workgroup of the SCOPE study. (2013). Cost analysis of the treatments for patients with advanced Parkinson's disease: SCOPE study. *Journal of Medical Economics*, *16*(2), 191-201.
- Volkman, J., Albanese, A., Antonini, A., Chaudhuri, K. R., Clarke, C. E., De Bie, R. M., ... & Nyholm, D. (2013). Selecting deep brain stimulation or infusion therapies in advanced Parkinson's disease: an evidence-based review. *Journal of Neurology*, *260*(11), 2701-2714.
- Westin, J., Nyholm, D., Pålhagen, S., Willows, T., Groth, T., Dougherty, M., & Karlsson, M. O. (2011). A pharmacokinetic-pharmacodynamic model for duodenal levodopa infusion. *Clinical Neuropharmacology*, *34*(2), 61-65.
- Willows, T., Dizdar, N., Nyholm, D., Widner, H., Groth, K., Larsson, J., ... & Kjellander, S. (2016). Telemedicine facilitates efficient and safe home titration of levodopa/carbidopa intestinal gel (LCIG) in patients with advanced Parkinson's disease. *Parkinsonism & Related Disorders*, *22*, e95

