

Aus der Klinik für Neurologie
Direktor: Prof. Dr. med. Lars Timmermann
Des Fachbereichs Medizin der Philipps-Universität Marburg

In Zusammenarbeit mit dem Oxford Parkinson's Disease Center

**Smartphone-basierte Untersuchung zu
Stimmcharakteristik, Körperbewegungen
und Geschicklichkeit bei Patienten mit REM-
Schlaf-Verhaltensstörung, Parkinson-
Krankheit und gesunden Kontrollpersonen**

INAUGURALDISSERTATION

zur Erlangung des Doktorgrades
des Fachbereichs Medizin der
Philipps-Universität Marburg

vorgelegt von

Moritz Beilharz

aus

München

Marburg, 2024

Angenommen vom Fachbereich Medizin
der Philipps-Universität Marburg
am: 10.01.2024

Gedruckt mit Genehmigung des Fachbereichs Medizin

Dekanin: Frau Prof. Dr. D. Hilfiker-Kleiner

Referent: Herr Prof. Dr. Dr. h.c. W. Oertel

1. Korreferent: Herr Prof. Dr. U. Koehler

Für meine Familie

Inhaltsverzeichnis

Abbildungsverzeichnis	IV
Tabellenverzeichnis	IV
Abkürzungsverzeichnis	V
1 Einleitung	1
2 Theoretischer Hintergrund	5
2.1 Die Parkinson-Krankheit (PK).....	5
2.1.1 Epidemiologie.....	6
2.1.2 Risikofaktoren.....	7
2.1.3 Ätiologie	8
2.2 Prodromalstadien der Parkinson-Krankheit.....	9
2.2.1 Vegetative Symptome.....	10
2.2.2 Beeinträchtigungen des Geruchssinns	10
2.2.3 Stimmveränderungen bei der PK.....	11
2.3 Die isolierte REM-Schlaf-Verhaltensstörung (RBD)	12
2.3.1 Epidemiologie.....	12
2.3.2 Ätiologie	14
2.3.3 Risikofaktoren.....	14
2.4 Aktuelle Ansätze in Therapie und Forschung.....	15
2.4.1 Selbstlernende Algorithmen und ihr Potenzial in der klinischen Forschung.....	16
3 Fragestellung.....	18
4 Material und Methoden	19
4.1 Finanzierung	19
4.2 Ein- und Ausschlusskriterien	19
4.2.1 Kohorte „RBD“	19
4.2.2 Kohorte „PK“.....	20
4.2.3 Kohorte „Gesunde Kontrollen“	20
4.2.4 Abbruchkriterien	20
4.3 Applikation	20
4.4 Technische Grundlagen der Testfunktionen	23
4.5 Smartphones.....	23
4.5.1 HTC Desire 310.....	23
4.5.2 Wiko Lenny 2	23
4.6 Probanden-Information	23

4.7	Vorgehen und Ablauf.....	24
4.7.1	Rekrutierung	24
4.7.2	Einweisung.....	24
4.7.3	Klinische Messung.....	24
4.7.4	Heimmessung.....	25
4.7.5	Rücksendung und Upload der generierten Daten	25
4.8	Auswertung der Daten	25
4.9	Statistische Auswertung der Ergebnisse	26
5	Ergebnisse	29
5.1	Demographische und klinische Charakteristik der Probanden	29
5.2	Testleistung.....	31
5.2.1	Testleistung nach LORO-Kreuzvalidierung und 10-fach-Kreuzvalidierung für die Unterscheidung zwischen RBD und Kontrollen.....	31
5.2.2	Herkunft der einflussreichsten Merkmale RBD vs. Kontrollen.....	33
5.2.3	Testleistung nach LORO-Kreuzvalidierung und 10-fach-Kreuzvalidierung für die Unterscheidung zwischen PK und Kontrollen	34
5.2.4	Herkunft der einflussreichsten Merkmale PK vs. Kontrollen.....	36
5.2.5	Testleistung nach LORO-Kreuzvalidierung und 10-fach-Kreuzvalidierung für die Unterscheidung RBD vs. PK	37
5.2.6	Herkunft der einflussreichsten Merkmale PK vs. RBD.....	39
5.3	Aussagekraft und Wichtigkeit der Merkmale	40
5.3.1	Aussagekraft der wichtigsten Merkmale RBD vs. Kontrollen.....	43
5.3.2	Aussagekraft zweier wichtiger Merkmale PK vs. Kontrollen	45
5.3.3	Aussagekraft der wichtigsten Merkmale RBD vs. PK.....	47
5.4	Anteile der Tests an den wichtigsten Merkmalen mit Aussagekraft zu allen Kohorten	49
6	Diskussion	51
6.1	Interpretation der Ergebnisse	51
6.2	Durchführbarkeit.....	55
6.2.1	Studienabbrecher und Compliance	55
6.2.2	Einfluss des Alters der Probanden	55
6.2.3	Probleme mit der Software und Hardware.....	56
6.3	Eigene Limitationen und Optimierungsmethoden	57
6.3.1	Einfluss des Geschlechts der Probanden.....	57
6.3.2	Limitation der Applikation und des Studiendesigns	58
6.3.3	Statistische und methodische Limitationen.....	60
6.4	Bedeutung der Ergebnisse und Fazit.....	61
6.5	Ausblick.....	62
7	Zusammenfassung.....	64
8	Summary.....	66
9	Literaturverzeichnis.....	67

10	Anhang	73
10.1	Informationsblatt zur Messungsdurchführung	73
10.2	Informationen und Aufklärung für Studienteilnehmer	77
10.3	Ethikantrag.....	82
10.4	Votum der Ethikkommission	110
	Verzeichnis der akademischen Lehrer und Lehrerinnen	112
	Danksagung.....	113

Abbildungsverzeichnis

Abbildung 1: „Random Forests“	27
Abbildung 2: Testleistung LORO RBD vs. Kontrollen	32
Abbildung 3: Testleistung 10-fach-Kreuzvalidierung RBD vs. Kontrollen.....	33
Abbildung 4: Unterscheidungsmerkmale RBD vs. Kontrollen.....	34
Abbildung 5: Testleistung LORO PK vs. Kontrollen	35
Abbildung 6: Testleistung 10-fach-Kreuzvalidierung PK vs. Kontrollen.....	36
Abbildung 7: Separationsmerkmale PK vs. Kontrollen	37
Abbildung 8: Testleistung LORO PK vs. RBD	38
Abbildung 9: Testleistung 10-fach-Kreuzvalidierung PK vs. RBD.....	39
Abbildung 10: Separationsmerkmale PK vs. RBD	40
Abbildung 11: Streudiagramm der zwei wertvollsten Merkmale in der Auswertung RBD vs. Kontrollen	43
Abbildung 12: Boxplots wertvollste Merkmale RBD vs. Kontrollen	45
Abbildung 13: Streudiagramm der zwei wertvollen Merkmale in der Auswertung PK vs. Kontrollen.....	46
Abbildung 14: Boxplots wertvolle Merkmale PK vs. Kontrollen.....	47
Abbildung 15: Streudiagramm der zwei wertvollsten Merkmale in der Auswertung RBD vs. PK	48
Abbildung 16: Boxplots wertvollste Merkmale RBD vs. PK	49
Abbildung 17: Zusammensetzung der Top 30 Merkmale insgesamt.....	50

Tabellenverzeichnis

Tabelle 1: Demographische Daten	29
Tabelle 2: Klinische Daten (Mittelwerte \pm Standardabweichung).....	29
Tabelle 3: ausgewertete Daten	30
Tabelle 4: Top 30 Merkmale der 10-fachen Kreuzvalidierung	42

Abkürzungsverzeichnis

AUC	Fläche unter der Kurve (engl. „Area under the Curve“)
BDI	Beck-Depressions-Inventar
DaTSCAN	Dopamin-Transporter-Szintigrafie (engl. „Dopamine Transporter Scan“)
DGN	Deutsche Gesellschaft für Neurologie
H&Y	Hoehn und Yahr Stadium
IMU	inertiale Messeinheit (engl. „Inertial Measurement Unit“)
KV	Kreuzvalidierung (engl. „Cross-Validation“)
LORO	„Leave-One-(Recording)-Out“-Kreuzvalidierung
MoCA	Engl.: „Montreal-Cognitive-Assessment“
NSAR	Nicht-Steroidale-Anti-Rheumatika
PDNMS-Skala	Engl.: „Parkinson-Disease-Non-Motor-Symptoms“-Skala
PK	Parkinson-Krankheit
RBD	REM-Schlaf-Verhaltensstörung (engl. „REM- Sleep Behavior Disorder“)
RBD-SQ	Engl.: „REM-Sleep-Behavior-Disorder-Screening-Questionnaire“
REM	Rasche Augenbewegung (engl. „Rapid Eye Movement“)
ROC	Operationscharakteristik eines Beobachters (engl. „Receiver operating characteristics“)
SD	Standardabweichung (engl.: „Standard deviation“)
SNpc	Substantia Nigra pars compacta
UPDRS	Engl. „Unified Parkinson’s Disease Rating Scale“
UPSIT	Engl. „University of Pennsylvania Smell Identification Test“

1 Einleitung

In dieser Arbeit wird die Studie der Philipps-Universität Marburg vorgestellt, die die Möglichkeiten, Schwierigkeiten und Erfolge der Applikation (Kurzform: App) „HopkinsPD“ in der Vermessung von Bewegungsstörungen, insbesondere der Parkinson-Krankheit (PK) und der REM-Schlaf-Verhaltensstörung (RBD), beleuchtet. In Fortsetzung der Arbeit der Universität Oxford beschäftigten wir uns mit einem eigenen Datensatz und möglichen zukünftigen Anwendungsbereichen sowie den Schwierigkeiten, die es auf dem Weg zur Marktreife noch zu überwinden gilt. Dabei lag unser Fokus auf der Übertragbarkeit der Ergebnisse unserer Auswertung mittels statistischen maschinellen Lernens („Machine-Learning“) auf den klinischen Alltag.

Die Anwendung digitaler Hilfsmittel zur Beurteilung des klinischen Zustandes im häuslichen Umfeld – „Home Assessment“ (engl. für Heim-Beurteilung) – ist ein Trend in der klinischen Forschung, dem großes Potential nachgesagt wird. Schon seit längerem besteht die Möglichkeit, Patienten¹ in ihrer häuslichen Umgebung objektiv zu untersuchen, ohne dass medizinisches Personal die Daten erheben muss. Die digitale Revolution und die Fortschritte in der Entwicklung von direkt am Körper getragenen Computersystemen, sog. „Wearables“, und Smartphones eröffnen allerdings eine nie dagewesene Zahl an neuen Untersuchungsansätzen. Sie werden heute von rund 74% der Deutschen über 14 Jahren genutzt und eine Ende des Booms wird laut Zahlen des Branchenverbands „Bitkom“ nicht erwartet (www.bitkom.org, 2016).

Smartphones erheben mittlerweile eine ganze Fülle an verschiedenen Gesundheitsdaten. Das tun sie entweder aktiv, wie zum Beispiel beim Joggen mit Puls- und Atemfrequenzmesser, oder passiv im Hintergrund als Schrittzähler. Inzwischen gibt es eine große Anzahl von Anbietern, die verschiedenste kommerzielle Fitness- und Gesundheits-Apps auf den Markt gebracht haben.

¹ Aus Gründen der besseren Lesbarkeit wird in dieser Arbeit das generische Maskulinum verwendet. Diese Formulierungen umfassen gleichermaßen männliche und weibliche Personen; alle sind damit selbstverständlich gleichberechtigt angesprochen.

Auch im medizinischen Bereich gibt es verschiedene Apps, die Patienten erlauben, ihre Gesundheit einzuschätzen oder Daten zur Auswertung für sich oder den Arzt zu sammeln. So bieten z.B. die Krankenkassen die Vivy-App als eine „digitale Gesundheitsassistentin“ an (<https://www.vivy.com/>), die z.B. medizinische Notfalldaten speichert, an Impfauffrischungen oder Medikamenteneinnahme erinnert, bei der Arztsuche hilft, Gesundheits-Checks anbietet oder die Möglichkeit hat, Röntgenbilder zu speichern und dem Arzt digital zu übermitteln. Ein anderes Beispiel ist die „SnoreLab“-App (<https://www.snorelab.com/>), die Schlaf- und Schnarchverhalten aufzeichnet, Schlafstatistiken erstellt, auswertet und Therapieempfehlungen gibt.

Die Entwicklung von medizinischen Applikationen wird in den letzten Jahren durch die technischen Fortschritte in der Computerentwicklung noch beflügelt. Ausgeklügelte Sensorelektronik, leistungsstarke Prozessoren und große Speicherkapazitäten auf kleinstem Raum machen Smartphones heute zu präzisen Messinstrumenten, die den Vorteil haben, dass sie die meisten Menschen täglich nutzen und zudem nah am Körper mit sich führen. Zusätzlich sind sie in der Lage, Daten kabellos über das Internet zu übertragen, und ermöglichen dem Untersucher die Auswertung seiner erhobenen Parameter, ohne vor Ort sein zu müssen. Diese Entwicklung bietet natürlich eine Vielzahl an attraktiven Untersuchungsmöglichkeiten, jedoch sollten auch Bedenken von Datenschützern berücksichtigt werden. Es handelt sich schließlich um sensible Gesundheitsdaten, die mit besonderer Vorsicht verarbeitet und ausgewertet werden müssen.

Besonders Bewegungsstörungen sind im Zusammenhang mit Smartphone-basierten Untersuchungen in den Fokus der Forscher gelangt, da sich hier großes Potenzial und neue Möglichkeiten zur Diagnostik aufgetan haben. Die Parkinson-Krankheit mit ihrem charakteristischen klinischen Bild mittels Bewegungssensorik zu erfassen, ist schon länger ein Ziel in der klinischen Forschung. Der Arbeitsgruppe um Dr. Siddharth Arora von der Universität Oxford gelang es mit diversen Smartphone-Tests, die auch in dieser Studie Verwendung fanden, mit einer Sensitivität von 96,2% und einer Spezifität von 96,9% Parkinsonpatienten von einer gesunden Kontrollgruppe zu unterscheiden (S. Arora et al., 2015). Das war ein wichtiger Schritt, da das idiopathische Parkinson-Syndrom (=Parkinson-Krankheit) als zweithäufigste neurodegenerative Erkrankung mit vielen verhältnismäßig jungen Betroffenen eine große gesellschaftliche Relevanz hat. Diese Pilot-Studie illustriert das Potenzial, welches in der Überwachung von Symptomen mittels Smartphones steckt.

Die folgende Studie der Arbeitsgruppe „prodromale Parkinson-Syndrome“ der Klinik für Neurologie der Universität Marburg leistet in der Weiterentwicklung des in der Pilot-Studie gewählten Ansatzes zur Untersuchung von Stimmcharakteristik, Geschicklichkeit und Körperbewegungen bei Patienten mit REM-Schlaf-Verhaltensstörung, Parkinson-Krankheit und einer gesunden Kontrollkohorte einen Beitrag zur tiefergehenden Analyse der Datenstruktur. Dabei sollte insbesondere der Weg hin zur Übertragbarkeit von Ergebnissen einer statistischen „Machine-Learning“-Analyse auf das klinische Bild der Patienten bereitet werden.

In Bezug auf die Therapie und die Unterstützung von Therapieentscheidungen, die in Zukunft mit Applikationen, wie der hier Untersuchten, getroffen bzw. unterstützt werden sollen, ist es notwendig, dass diese Programme auch Abstufungen in der Symptomlast erkennen und objektivieren können. Die Hoffnung dabei ist, dass sie dies in Zukunft mindestens genauso gut oder besser als der Arzt in seiner Praxis können. Denn bei ähnlich validen Ergebnissen würden Applikationen kontinuierlicher, öfter und schneller untersuchen, und damit dem behandelnden Arzt einen gesamtheitlicheren Blick auf Symptomlast und Symptomfluktuation eröffnen. Die sensible Sensorik in modernen Smartphones könnte es in Zukunft ermöglichen, Symptome wie Tremor oder Ganginstabilitäten früher und präziser zu erkennen und zuzuordnen. Gerade im Hinblick auf in Zukunft hoffentlich verfügbare neuroprotektive Therapien bei der Parkinson-Krankheit ist eine frühe und präzise Diagnose essentiell.

Apps, die in der Untersuchung von Patienten mit Bewegungsstörungen eingesetzt werden, arbeiten in der Regel mit selbstlernenden Algorithmen, die extrem potent darin sind, Muster zu erkennen und daraus Vorhersagen zu treffen, wie z.B. die Zugehörigkeit zu einer Patientengruppe. Allerdings lassen sich aus ihnen keine Erklärungen ableiten, die zu einem umfassenderen Verständnis der Datenstruktur beitragen würden. Das verhindert den Vergleich mit bestehendem Wissen und nimmt somit einen der wichtigsten Validierungsparameter aus dem Spiel (Kubota, Chen, & Little, 2016).

Und dies ist nur eine der Schwierigkeiten, die es durch vorsichtiges und kleinschrittiges Vorgehen zu überwinden gilt. Unter anderem stellt die Qualität der Daten, die eben nicht mehr unter klinischer Aufsicht erhoben werden, ein Problem dar. Verknüpft man bei Parkinsonpatienten die erhobenen Daten zu den motorischen Symptomen etwa mit der Medikamenteneinnahme, so ist der Zeitpunkt der Messung in Bezug auf die Einnahme der Tabletten essentiell, um Fluktuationen in der Medikamentenwirkung und dadurch auch

Fluktuationen in der Testleistung abschätzen zu können. Versäumt ein Parkinsonpatient sein Medikament pünktlich zu nehmen oder erledigt die Tests nicht pünktlich, ergibt sich daraus möglicherweise ein inkongruentes Bild in den erhobenen Daten, das wiederum extrem schwierig zu entschlüsseln ist.

Der Weg zur Marktreife einer entsprechenden Applikation ist dementsprechend lang und die Hürden, die es zu überwinden gilt sind hoch, um das Potenzial dieser Entwicklungen zu nutzen. Ein potentieller Vorteil liegt darin, dass die digitalen Daten über „Cloud-dienste“ für viele verschiedene Entwickler zugänglich gemacht werden können, und ein Nischenprodukt so trotzdem die ausreichende Datenmenge zur Validierung erreicht. Dies würde auch die kontinuierliche Weiterentwicklung und Verfeinerung solcher Diagnosewerkzeuge erlauben, wobei die Frage des Datenschutzes und der Finanzierung geklärt werden muss, da für private Investoren der Unterhalt eines solchen Programms keine ausreichenden finanziellen Anreize gewährt (Kubota et al., 2016).

Technisierung und Objektivierung ermöglichen in der Medizin seit Jahrzehnten Fortschritte, die sich vor hundert Jahren niemand hätte träumen lassen. Insbesondere im relativ jungen Feld der digitalen Hilfsmittel zur häuslichen Beurteilung („Home Assessments“), das erst durch die breite Verfügbarkeit von Technik in den letzten Jahrzehnten entstanden ist, schlummert großes Potenzial, dass es noch zu erforschen gilt. In der Medizin hat sich in den letzten 20 Jahren viel getan. In verschiedensten Disziplinen sind Umwälzungen des Status quo im Gange, die unter anderem von der Digitalisierung und Technisierung getragen sind. Eben auch in der Parkinsonforschung hält die Digitalisierung in Bereichen Einzug, die vor 10 Jahren noch einzig von einem klinischen Untersucher erfasst werden konnten. Dazu will die folgende Studie, die an der Klinik für Neurologie der Philipps-Universität Marburg in Zusammenarbeit mit dem Oxford Parkinson's Disease Center durchgeführt wurde, nun einen Beitrag leisten.

2 Theoretischer Hintergrund

Die Parkinson-Krankheit ist die zweithäufigste neurodegenerative Erkrankung nach der Alzheimer-Demenz (Dorsey et al., 2007; Fargo & Bleiler, 2014) und die häufigste Bewegungsstörung des Menschen (Tysnes & Storstein, 2017). Insbesondere mit dem Altern der Bevölkerung wird erwartet, dass die Inzidenz weiter steigen wird (Von Campenhausen et al., 2005). Dementsprechend wird substanzieller Aufwand in Forschung und Therapie betrieben, um die komplexe Krankheit in Symptomen und Krankheitsprogression zu mildern, wobei nach wie vor keine Therapie bekannt ist, welche das Fortschreiten der Krankheit verlangsamen kann.

Mit zunehmendem Verständnis der Krankheit und dem Fortschritt in der Entwicklung neuer Medikamente in den letzten 20 Jahren, rücken aber auch Therapien, die den Krankheitsverlauf positiv beeinflussen könnten, in greifbarere Nähe (Levin & Oertel, 2021). Die Entwicklung krankheitsmodifizierender Therapien bedarf diverser Tests, die Parkinsonsymptome spezifisch erfassen und ein Fortschreiten der Erkrankung über relativ kurze Zeitintervalle detektieren können. In einer Pilotstudie an der Universität Oxford gelang es mittels Smartphones mit einer Spezifität und Sensitivität >95% Parkinsonpatienten von gesunden Kontrollpersonen zu unterscheiden (Arora et al., 2015). Dabei wurde mit Hilfe von statistischem „Machine-Learning“ eine Zuordnungsentscheidung zu einer der zwei untersuchten Gruppen getroffen.

Ein wesentlicher Aspekt der Untersuchung der Patienten mittels Smartphone-integrierter Trägheitsmessgeräte = Inertial-Measurement-Units (IMU) ist die geringere Anfälligkeit der Tests gegenüber einer Inter- und Intra-Prüfer Inkohärenz. Dieses Maß an Objektivität ist mit einer klassischen klinischen Untersuchung nicht zu erreichen, da zu oft verschiedene Ärzte denselben Patienten unterschiedlich beurteilen. Deshalb schafft die Smartphone-basierte Untersuchung eine einheitliche solide Datengrundlage, auf der Therapien in Zukunft beurteilt werden könnten. Inwieweit diese IMUs auch bei Prodromalstadien der Parkinson-Krankheit sinnvoll einsetzbar sind, soll unter anderem im Folgenden untersucht werden.

2.1 Die Parkinson-Krankheit (PK)

Parkinson-Syndrome werden unterteilt in das idiopathische Parkinson-Syndrom, genetische Formen des Parkinson-Syndroms, atypische Parkinson-Syndrome und sekundäre

Parkinson-Syndrome (DGN, 2016). Im Folgenden soll vor allem das idiopathische Parkinson-Syndrom betrachtet werden, welches entsprechend der klinischen Präsentation in drei verschiedene Verlaufsformen unterteilt wird. Man unterscheidet hierbei den akinetisch-rigiden Typ, den Äquivalenz-Typ und den Tremordominanz-Typ. Patienten mit einem Parkinson-Syndrom vom Tremordominanz-Typ weisen eine langsamere Krankheitsprogression auf als Patienten mit einem akinetisch-rigiden Typ (Oertel, Deuschl & Poewe, 2012).

Für die Diagnose des idiopathischen Parkinson-Syndroms existieren in der Literatur mehrere Empfehlungen, von welchen die „UK Parkinson’s Disease Society Brain Bank Clinical Diagnostic Criteria“ die geläufigsten sind (Gibb, 1988). Aktueller wissenschaftlicher Konsens ist, dass der pathologische Prozess, der schlussendlich in einem manifesten Parkinson-Syndrom resultiert, schon Jahre vorher beginnt (Hawkes, Tredici & Braak, 2010). In dieser prodromalen Phase wurden eine Reihe von nicht-motorischen Symptomen beobachtet, deren Auftreten das Risiko, an einem idiopathischen Parkinson-Syndrom oder einer anderen neurodegenerativen Erkrankung aus dem Formenkreis der atypischen Parkinson-Syndrome zu erkranken, massiv erhöht. Riechstörung, RBD, autonome Dysfunktionen, wie Obstipation oder orthostatische Dysregulation, und Depression sind dabei die häufigsten nicht-motorischen Symptome, die mit einer direkten Erhöhung des Parkinsonrisikos einhergehen (Postuma et al., 2019).

2.1.1 Epidemiologie

In Europa sind einer Schätzung nach ca. 1,2 Millionen Menschen an der PK erkrankt. In Deutschland beläuft sich die Patientenzahl noch auf ca. 220 000 Parkinsonpatienten. Hierbei ist das Verhältnis der Geschlechter ungleich zu Lasten der Männer verteilt, die ca. 1,46 mal so häufig an Parkinson erkranken (Gustavsson et al., 2011). Auf Grund des demographischen Wandels in den Industrienationen wird sich die Patientenlast jedoch bis zum Jahr 2030 verdoppeln, womit auch die finanziellen Belastungen der Gesundheitssysteme weiter steigen werden (Dorsey et al., 2007). Die Prävalenz wird dabei in der Literatur mit 108-257/100 000 Einwohnern und die Inzidenz auf 11-19/100 000 Einwohnern und Jahr beziffert, wobei Inzidenz und Prävalenz in der Bevölkerung mit steigendem Alter zunehmen (Von Campenhausen et al., 2005).

2.1.2 Risikofaktoren

Das Risiko, an einem Parkinson-Syndrom zu erkranken, ist nicht in allen Teilen der Welt gleich hoch. Die Prävalenz der Parkinson-Krankheit ist in Westeuropa, Nord- und Südamerika am höchsten (Bauso et al., 2012; Strickland & Bertoni, 2004; Von Campenhausen et al., 2005). Hinzu kommt, dass in den USA die Inzidenz für Hispanoamerikaner am höchsten ist, gefolgt von Kaukasiern, Menschen asiatischer Herkunft und Afroamerikanern (Van Den Eeden, 2003). Inzidenz und Prävalenz der Parkinson-Erkrankung steigen fast exponentiell mit zunehmendem Alter. Die Prävalenz ist in der Gruppe der über 80-Jährigen am höchsten (Pringsheim, Jette, Frolkis, & Steeves, 2014) Die höchste Inzidenz findet sich zwischen 85 und 89 Jahren (Driver, Logroscino, Gaziano, & Kurth, 2009).

Weiterhin konnten in einer Metaanalyse eine Vielzahl heterogener Einflussfaktoren auf das Risiko, an einer Parkinson-Krankheit zu erkranken, identifiziert werden. Familiäre Belastung scheint dabei das Erkrankungsrisiko am stärksten zu erhöhen, gefolgt von Obstipation, affektiven Störungen, Pestizid-Exposition, Kopf-Verletzungen, Leben in ländlichen Regionen, Verwendung von Beta-Blockern, landwirtschaftlichen Tätigkeiten und dem Trinken von Brunnenwasser in der Vorgeschichte (Noyce et al., 2012). Als protektive Faktoren konnten Zigaretten-, Kaffeekonsum, sowie vorbestehender arterieller Hypertonus, NSAR-Einnahme, Einnahme von bestimmten Calcium-Kanal-Blockern und Alkoholkonsum identifiziert werden (Noyce et al., 2012). Personen, die Tabak rauchen, scheinen ein besonders geringes Risiko zu haben, an Parkinson zu erkranken. Ob dabei Nikotin, wie lange vermutet, tatsächlich einen positiven Einfluss auf die Entwicklung der Krankheit hat, wird von einer neueren Fall-Kontroll-Studie allerdings in Frage gestellt (Ritz, Lee, Lassen, & Arah, 2014). Diese konnte nämlich zeigen, dass Parkinsonpatienten leichter das Rauchen aufgeben konnten, was eine verminderte Nikotinsensibilität in der Prodromalphase der Parkinson-Krankheit nahelegt (Ritz et al., 2014).

Die starke Häufung bzw. Risikoerhöhung bei familiärer Vorbelastung mit Parkinson- oder Tremor-Erkrankungen legt den Schluss nahe, dass es eine starke genetische Komponente gibt. In einer Metaanalyse konnten 24 Gen-Loci identifiziert werden, die eine klinisch signifikante Assoziation zum Krankheitsrisiko aufwiesen (Nalls et al., 2014).

Schon die unterschiedlichen Risikofaktoren zeigen, dass der Parkinson-Krankheit ein komplexes Zusammenspiel verschiedenster ätiologischer Faktoren zugrunde liegt, welches noch nicht endgültig verstanden ist.

2.1.3 Ätiologie

Die verschiedenen zum Teil noch unverstandenen Ursachen der Parkinson-Krankheit bedingen schlussendlich einen Verlust dopaminergener Neurone in der Substantia Nigra Pars Compacta, die in das Striatum projizieren. Dieser Verlust der Neurone wird seit Ehringer und Hornykiewicz als ursächlich für die typischen motorischen Symptome wie Bradykinese verbunden mit Rigor, Tremor und posturaler Instabilität angesehen (Ehringer & Hornykiewicz, 1960). Die wiederum als ursächlich für den Untergang der dopaminergen Neurone angesehenen Pathologien, finden sich allerdings schon viel früher bei betroffenen Patienten. Das Braak-Modell integriert diese Erkenntnisse zu einem Gesamtkonzept, mit dem sich verschiedene Stadien der Erkrankung unterscheiden lassen (Braak et al., 2003). Prodromalen Parkinsonsymptomen konnten dadurch erklärende Pathologien zugeordnet werden.

Die ersten pathologischen Veränderungen im Sinne von alpha-Synuclein-Aggregate enthaltenden Einschlusskörperchen (Lewy-Körperchen) finden sich schon einige Zeit vor dem ersten Auftreten somatomotorischer Dysfunktionen. Dabei unterscheidet sich nur die räumliche Verbreitung der Lewy-Körperchen zwischen symptomatischen und nicht-symptomatischen Fällen, hingegen nicht deren Zusammensetzung (Braak et al., 2003). Das Braak-Modell besagt, dass sich in Stadium 1 der Erkrankung die ersten alpha-Synuclein-Aggregat-positiven Einschlüsse vor allem in den Kernen der Hirnnerven X und XI, sowie dem Bulbus olfactorius finden. In Stadium 2 werden zusätzlich die kaudalen Raphe-Kerne, der Nucleus gigantocellularis reticularis, sowie der Nucleus coeruleus und subcoeruleus involviert. Stadium 1 und 2 scheinen hierbei die nicht motorischen Symptome der Prodromalphase zu erklären (mit autonomen Dysfunktionen oder Verlust des Geruchsempfindens) (Braak et al., 2003).

Stadium 3 und 4 markieren den Beginn und das Fortschreiten der motorischen Symptomatik. Lewy-Körperchen finden sich ab Stadium 3 auch im Mesencephalon und dabei vor allem in der Pars compacta der Substantia nigra, wobei die typische Depigmentierung, die in späten Stadien erkennbar wird, noch nicht makroskopisch sichtbar ist. Diese zeigt sich ab Stadium 4, wobei die neuropathologischen Veränderungen weiter aufsteigen und nun auch das basale Prosencephalon und den Allocortex betreffen (Braak et al., 2003).

Ab Stadium 5 ist der Neocortex in die pathologischen Prozesse involviert. Stadium 5 und 6 markieren die späten Stadien der Parkinson-Krankheit und es finden sich stark fortgeschrittene neuropathologische Veränderungen in allen vorher genannten Abschnitten des

Gehirns. Ab Stadium 6 ist praktisch der gesamte Neocortex von Parkinson assoziierten Pathologien durchzogen. (Braak et al., 2003).

Das Aufsteigen der spezifischen Pathologien und die damit offenbar zusammenhängende klinische Verschlechterung der Patienten scheint eine ähnliche Pathogenese wie bei Prionen-Erkrankungen wahrscheinlich zu machen. Jedoch scheint es einige Variabilität zwischen den verschiedenen Fällen von Parkinson zu geben, sodass keinesfalls alle Erkrankungen und Pathologien über eine aufsteigende Transmission von alpha-Synuclein-Aggregaten erklärt werden können (Visanji, Brooks, Hazrati, & Lang, 2013) (Review).

Zudem ist auch eine Beteiligung von entzündlichen Prozessen an dem Krankheitsgeschehen anzunehmen. Die Art und Weise, wie neuroinflammatorische Prozesse in die Pathogenese der Parkinson-Krankheit eingreifen, ist aber noch nicht vollständig verstanden. Für die regelmäßige Einnahme von NSAR konnte eine negative Assoziation zur PK nachgewiesen werden (Noyce et al., 2012). Daher scheint der Schluss naheliegend, dass inflammatorische Prozesse den Untergang von Neuronen begünstigen, die relevanten Mechanismen hierfür müssen allerdings noch geklärt werden (Kalia & Lang, 2015) (Review).

2.2 Prodromalstadien der Parkinson-Krankheit

Mit der Entwicklung der Braak-Stadien begann eine neue Ära in der Betrachtungsweise der Parkinson-Krankheit. Die Feststellung, dass die Neurodegeneration beginnt, bevor es zu motorischen Symptomen kommt, hat die Suche nach Prodromalstadien der Parkinson-Krankheit befeuert. Das spezifische Muster der neuronalen alpha-Synuclein-Ablagerung legt nahe, dass einige Symptome der Parkinson-Krankheit direkt auf Neurodegenerationen zurückzuführen sind, die vor dem Auftreten der ersten motorischen Symptome nachweisbar sind (Braak et al., 2003). Die Tatsache, dass die Parkinson-Krankheit nicht in der Substantia Nigra und mit dem Auftreten der ersten motorischen Symptome beginnt, hat dazu geführt, dass in der jüngeren Vergangenheit zahlreiche nicht motorische, vermutlich prodromale Symptomkomplexe untersucht wurden, die ein hohes Potential haben könnten, die PK zu erkennen, bevor die eigentlichen motorischen Symptome augenfällig werden. Auf die vielversprechendsten dieser Symptome soll im Folgenden eingegangen werden.

2.2.1 Vegetative Symptome

Autonome Dysregulation ist ein wichtiger Symptomkomplex, der offenbar in einem frühen Stadium der Parkinson-Krankheit auftritt (Ziemssen et al., 2011). Das Problem der autonomen Dysregulation ist ihre fehlende Spezifität für neurodegenerative Erkrankungen. Die Volkskrankheit Diabetes, um nur ein Beispiel zu nennen, geht in fortgeschrittenen Stadien regelhaft mit einer Dysfunktion der autonomen Nerven einher und verursacht dadurch eine Vielzahl verschiedener Beschwerden, unter anderem Obstipation und orthostatische Dysregulation (Strian & Haslbeck, 1986). Auch Patienten mit Parkinson haben eine höhere Wahrscheinlichkeit, in ihrer Vorgeschichte an Obstipation gelitten zu haben (Postuma et al., 2012). Insbesondere gastrointestinale Beschwerden werden von einer schier endlosen Liste an Krankheiten verursacht, insofern ist abzusehen, dass das Erkennen einer prodromalen Parkinson-Krankheit aufgrund von gastrointestinalen Beschwerden extrem schwierig ist. Damit reiht sich die autonome Dysregulation in eine Liste von Symptomkomplexen, die bei Betrachtung des Gesamtbildes auf eine Parkinson-Krankheit hindeuten, aber häufig übersehen werden, da Veränderungen subtil und wenig spezifisch sind und meist für sich alleine stehend betrachtet werden. Gerade deshalb ist ein interdisziplinärer Ansatz äußerst wichtig, zumal Symptome der autonomen Dysregulation mehr Einfluss auf die Lebensqualität haben können als die motorischen Symptome (Ziemssen et al., 2011).

2.2.2 Beeinträchtigungen des Geruchssinns

Über 70% der Patienten mit einer vermuteten Parkinson-Krankheit haben im Vergleich zu einer gleichaltrigen Kohorte eine signifikante Beeinträchtigung ihres Geruchssinns (Hawkes et al., 1997). Weitergehend wird vermutet, dass die verbleibenden 20-30% in Wahrheit keine Parkinson-Krankheit haben (Postuma et al., 2012). Insbesondere da auch Patienten mit Lewy-Körperchen-Demenz oder selten auch mit Multisystematrophie eine Beeinträchtigung des Geruchssinns zeigen, scheinen Fehldiagnosen in der untersuchten Kohorte nicht unwahrscheinlich (Hawkes, 2003; Olichney et al., 2005; Shah, Muhammed, Findley, & Hawkes, 2008). Da bei progressiver supranukleärer Blickparese, corticobasaler Degeneration und medikamentös induzierten Parkinsonsyndromen Beeinträchtigungen im Geruchsempfinden eher ungewöhnlich sind, sollte eine fehlende Anosmie zum Überdenken der Diagnose führen (Postuma et al., 2012).

Verschiedenste Studien haben gezeigt, dass eine Riechstörung häufig im Prodromalstadium der Parkinson-Krankheit zu finden sein kann (Jennings et al., 2017; Miglis et al., 2021). Ein Problem beim Erfassen einer Riechstörung liegt in den uneinheitlichen Methoden der Analyse, sodass unterschiedlich sensitive Tests auch zu unterschiedlichen Ergebnissen führen und damit auch eine präzise Aussage hinsichtlich der Dauer der Prodromalphase erschweren (Postuma et al., 2012).

Da eine rein anamnestische Erhebung der Geruchswahrnehmung weder ausreichend sensitiv noch spezifisch ist, muss die Geruchswahrnehmung in einem kontrollierten Umfeld unter reproduzierbaren Bedingungen gemessen werden zum Beispiel mittels des „University of Pennsylvania Smell Identification Tests“ (UPSIT) oder der „Sniffin‘ Sticks“ (Burghart Messtechnik, Deutschland) (Doty et al., 1984; Hummel et al., 1997). Eine kontrollierte Erhebung mit einem UPSIT ist ähnlich sensitiv in der Diagnose der Parkinson-Krankheit wie ein Dopamine Transporter Scan (DaTSCAN), jedoch ist keine der beiden Untersuchungen spezifisch hierfür (Deeb et al., 2010). In Zusammenschau kann eine Analyse des Geruchssinns einen frühen Hinweis auf das Vorliegen einer Prodromalphase der Parkinson-Krankheit geben, sie ist allerdings nur gemeinsam mit anderen Untersuchungen sinnvoll einzusetzen (Miglis et al., 2021).

2.2.3 Stimmveränderungen bei der PK

Eine eingeschränkte Frequenzvariabilität, eine leise und heisere Stimme gehören zu den typischen Veränderungen bei nicht therapierten Parkinson-Kranken, wobei die Intensität mit einem späteren Krankheitsstadium korreliert (Holmes, Oates, Phyland, & Hughes, 2000). Ähnliche Symptome können auch schon bei Patienten mit isolierter REM-Schlaf-Verhaltensstörung (RBD) nachgewiesen werden, sodass man Stimmveränderungen als einen vielversprechenden Screening-Parameter ansieht. Mit einer Sensitivität von 96% und einer Spezifität von 79% konnte eine Analyse der vier prominentesten Stimmveränderungen bei RBD eine Kontrollgruppe von einer RBD-Kohorte unterscheiden (Rusz et al., 2016). Eine weitergehende Analyse mit größeren Kohorten konnte diverse Veränderungen von Stimme und Sprache dokumentieren und präsentierte gleichzeitig einen Ansatz der vollautomatischen Analyse (Hlavnička et al., 2017).

Die automatisierte Analyse von Sprache und Stimme ist dabei ein sehr vielversprechender Ansatz, sowohl um Patienten mit RBD zu erkennen als auch um ein Fortschreiten der Erkrankung und ein Einsetzen der ersten diskreten motorischen Symptome zu detektieren

und vorherzusagen. Insbesondere die Analyse von zusammenhängender Sprache in der häuslichen Umgebung, die zwar einigermaßen frei von Störgeräuschen sein sollte, aber keinen aufwändigen Versuchsaufbau benötigt, ist aufgrund ihrer Tauglichkeit zum Screening einer breiten Masse besonders interessant. Kombiniert man die Analyse der Stimme nun mit anderen Tests, zum Beispiel der Motorik, lassen sich Sensitivität und Spezifität noch steigern. Ziel dieser Bemühungen sollte nicht alleine eine Zuordnung zu PK, RBD oder gesund sein, denn langfristig ist auch das Monitoring der Symptomlast und der Krankheitsprogression entscheidend, um die Therapiesteuerung zu optimieren oder auch krankheitsmodifizierende Medikamente zu testen.

2.3 Die isolierte REM-Schlaf-Verhaltensstörung (RBD)

Die isolierte REM-Schlaf-Verhaltensstörung ist eine Parasomnie, die durch den Verlust der Atonie während der REM-Schlaf-Phasen definiert ist (American Academy of Sleep Medicine, 2005). Betroffene haben häufig lebhaft bis gewalttätige Träume, die sie aufgrund der fehlenden Atonie in der REM-Schlaf-Phase derart ausleben, dass sie Gefahr laufen, sich oder ihre Lebenspartner zum Teil ernsthaft zu verletzen (Schenck & Mahowald, 2002).

Um die Diagnose einer REM-Schlaf-Verhaltensstörung zu stellen, bedarf es wiederholter, typischer Episoden mit Vokalisationen und komplexen Bewegungen aus dem REM-Schlaf und der Dokumentation des REM-Schlafs ohne Atonie mittels Video-Polysomnographie (PSG) (International Classification of Sleep Disorders, 3rd Edition, AASM, 2014). Ohne PSG kann nur die Verdachtsdiagnose einer RBD bei klaren und eindeutigen klinischen Hinweisen gestellt werden. Zur Diagnose einer isolierten REM-Schlaf-Verhaltensstörung darf diese nicht durch andere Erkrankungen, wie z.B. Narkolepsie oder Medikamenteneinnahme, erklärt sein (Sateia, 2014).

Zur Therapie der unerwünschten Nachtaktivität werden vor allem Clonazepam und Melatonin eingesetzt (Dauvilliers et al., 2018).

2.3.1 Epidemiologie

Betrachtet man die Epidemiologie der isolierten REM-Schlaf-Verhaltensstörung, so fallen zwei Dinge besonders auf: Der typische RBD-Patient ist männlich und zwischen 50-65 Jahre alt. 87% der 93 untersuchten Patienten in der Studie der Mayo Clinic waren

männlich (Olson, Boeve, & Silber, 2000). Diese Daten konnten auch Schenck und Mahowald mit 87,5% männlichen von 96 untersuchten Patienten bestätigen (Schenck & Mahowald, 2002). Das mittlere Alter bei Beginn der Erkrankung war 52,4 Jahre (Schenck, Hurwitz, & Mahowald, 1993), und 60,9 Jahre (Olson et al., 2000). Diese Ergebnisse decken sich mit weiteren Untersuchungen u.a. aus Straßburg (Sforza, Krieger, & Petiau, 1997). Allerdings geben die Autoren zu bedenken, dass es durchaus möglich ist, dass mildere Formen der RBD, nämlich ohne aggressive Träume bei Frauen, bestehen könnten, und es deshalb eine Dunkelziffer auf Grund unzureichender Sensitivität der Diagnosekriterien gibt (Schenck & Mahowald, 2002).

Die enge Assoziation zwischen PK und RBD zeigt sich auch, wenn man einen Blick auf die Prävalenz von (in dem Fall sekundärer) RBD bei an Parkinson erkrankten Patienten wirft. Eine chinesische Metaanalyse konnte zeigen, dass 42,3% aller Patienten mit einem idiopathischen Parkinson-Syndrom auch an einer (sekundären) REM-Schlaf-Verhaltensstörung leiden. Eine unterschiedliche Handhabung der Diagnosekriterien für RBD und eine Heterogenität bezüglich des Alters der in den Studien untersuchten Personen, sowie Unterschiede in der Erkrankungsdauer, waren dabei die Hauptgründe für Variationen in der Prävalenz in einigen der untersuchten Studien (Zhang, Sun, Wang, Tang, & Xie, 2017).

Umgekehrt konnte eine Metaanalyse zeigen, dass Patienten mit RBD ein Risiko von über 90% haben, nach 14 Jahren an einer neurodegenerativen Erkrankung zu erkranken. Die mittleren Konversionsraten lagen nach 5 Jahren bei 33,6%, nach 10,5 Jahren bei 82,4% und nach 14 Jahren bei 96,6%. Die meisten Probanden (43%) entwickelten eine Parkinson-Krankheit, gefolgt von 25%, die an einer Lewy-Körperchen-Demenz erkrankten (Galbiati, Verga, Giora, Zucconi, & Ferini-Strambi, 2019). Die bisher umfangreichste multizentrische Studie hierzu konnte zeigen, dass die mittlere Konversionsrate bei RBD pro Jahr bei 6,25% liegt. 56,5% der Patienten mit RBD, die im Verlauf konvertierten, entwickelten eine Parkinson-Krankheit, wobei hiervon im Verlauf bei 4,5% der Verdacht auf eine Multisystematrophie bestand. 43,5% der Patienten entwickelten eine Lewy-Körperchen-Demenz (Postuma et al., 2019).

2.3.2 Ätiologie

Die enge Assoziation von alpha-Synucleinopathien und RBD legt nahe, dass der zugrundeliegende Pathomechanismus entweder der gleiche ist oder aber zumindest die gleichen cerebralen Strukturen betroffen sind bzw. zerstört werden.

Schlaf scheint im Gehirn durch die kombinierte Aktivität eines komplexen verteilten Netzwerks mit abgegrenzten Schlaf-Kreisläufen initiiert und aufrechterhalten zu werden. Dabei spielen insbesondere Zellpopulationen in der Area praeoptica und GABAerge/glycinerge Neurone der parafazialen Zone (PZ) eine wichtige Rolle (Anaclet et al., 2012).

REM-Schlaf insbesondere ist durch die Aktivität und Interaktion einer Reihe von weiteren Hirnregionen in Hypothalamus, Vorderhirn und Hirnstamm gekennzeichnet. Eine Region von besonderem Interesse ist hierbei der Nucleus sublaterodorsalis, weil hier die Neurone vermutet werden, die Atonie und kortikale Aktivierung verursachen (Fragne, Torontali, Snow, & Peever, 2015). Es konnte eine hohe Vulnerabilität der im REM-Schlaf involvierten Regionen gegenüber alpha-Synuclein- und Tau-Pathologien gezeigt werden (Dugger, Boeve, et al., 2012; Dugger, Murray, et al., 2012). Im Tiermodell konnten jedoch einige Hirnregionen identifiziert werden, deren Fehlfunktion einen Verlust der Atonie während der REM-Schlaf-Phasen bedingen. Eine Dysfunktion des großzelligen medialen Anteils der *Formatio Reticularis*, des sublaterodorsalen Nucleus und *Peri-Locus Coeruleus* scheint zu REM-Schlaf ohne Atonie zu führen (Krenzer et al., 2011). Zum Vollbild der REM-Schlaf-Verhaltensstörung kommt es bei fortschreitender Degeneration dieser und möglicherweise zusätzlicher Areale (Boeve, 2013). Spezifischer konnte gezeigt werden, dass es bei Ratten zu Verlust von Atonie während des REM-Schlafs kommt, wenn man GABAerge/glycinerge-Neurone der ventromedialen Medulla gentechnisch inaktiviert (Valencia Garcia et al., 2018). Im REM-Schlaf werden diese inhibitorischen GABAergen/glycinergen Neurone von glutamatergen Neuronen des Nucleus sublaterodorsalis aktiviert und hemmen so die spinalen Motoneurone (Valencia Garcia et al., 2017). Kongruent zum Braak-Modell kann dies erklären, warum RBD der Parkinson-Krankheit vorausgehen kann.

2.3.3 Risikofaktoren

Über 80% der Patienten mit isolierter RBD entwickeln im Laufe ihres Lebens eine Parkinson-Krankheit, Lewy-Körper-Demenz oder sehr selten Multisystematrophie (sogenannte alpha-Synucleinopathien) (Schenck, Boeve, & Mahowald, 2013). Damit ist die

isolierte REM-Schlaf-Verhaltensstörung ein enormer Risikofaktor für die Entwicklung einer alpha-Synukleinopathie und stellt ein sehr spezifisches Prodromalstadium dar (Postuma et al., 2012). Deshalb sind die Risikofaktoren für eine Parkinson-Krankheit und RBD auch weitgehend ähnlich. In einer Multicenter Studie konnte gezeigt werden, dass Patienten mit RBD öfter Raucher sind, im Schnitt weniger Jahre institutionalisierter Schulbildung absolviert haben, öfter in der Landwirtschaft arbeiten und öfter Pestiziden ausgesetzt waren als die Kontrollgruppe. Außerdem hatten sie öfter in der Vergangenheit eine Kopfverletzung (Postuma et al., 2012). Die wichtigsten klinischen Risikofaktoren für eine Phänokonversion von RBD zu einer der oben genannten alpha-Synukleinopathien sind hierbei in absteigender Reihenfolge motorische Auffälligkeiten in der standardisierten Untersuchung, Riechstörung, kognitive Auffälligkeiten in der neuropsychologischen Testung, erektile Dysfunktion, ein pathologischer DaTSCAN, eingeschränktes Farbsehen, Obstipation, sowie das Alter (Postuma et al., 2019).

2.4 Aktuelle Ansätze in Therapie und Forschung

Die große Hoffnung der Parkinson-Forschung ist es, ein den Krankheitsverlauf beeinflussendes Medikament zu entwickeln. Mit den bisher zur Verfügung stehenden Medikamenten kann die Symptomlast für viele Jahre sehr gut kontrolliert werden und für Parkinsonpatienten kann noch viele Jahre nach der Diagnosestellung eine hohe Lebensqualität erhalten werden. Allerdings ist die Aussicht, dass die Krankheit unaufhaltsam fortschreitet, für die Betroffenen sehr belastend. Die Entwicklung von krankheitsmodifizierenden Therapien (engl. „Disease modifying Drugs“) hat daher in der Forschung den höchsten Stellenwert. Das jedoch gestaltet sich unter anderem deshalb als schwierig und teuer, da die Medikamente in der klinischen Erprobung möglicherweise über einen langen Zeitraum gegeben werden müssten, um subtile Verbesserungen bzw. ein Verlangsamen der Krankheitsprogression mit den gängigen klinischen Untersuchungstechniken überhaupt nachzuweisen. Noch dazu sind diese Tests untersucherabhängig und eine objektive Beurteilung des Gesamtbildes ist schwierig, wenn eine Untersuchung des Patienten immer nur sporadisch und nicht kontinuierlich erfolgen kann.

2.4.1 Selbstlernende Algorithmen und ihr Potenzial in der klinischen Forschung

An dieser Stelle kann digitalen Untersuchungsmöglichkeiten eine wichtige Aufgabe zukommen. Sie ermöglichen unter Umständen eine kontinuierlichere und objektivere Analyse bzw. Untersuchung des Patienten. Die Hoffnung ist dabei, dass sie subtilere Veränderungen der Krankheitsprogression schneller entdecken können als der Arzt und dass sie dabei Personal und Kosten sparen bei vergleichbarer oder besserer Testleistung (Kubota et al., 2016).

Die Daten, die auch in dieser Studie mit Hilfe von Smartphones erhoben wurden, sind allerdings nicht leicht zu verarbeiten. Ein ambivalenter Aspekt in der Natur eines Datensatzes, der mit Hilfe von Sensorelektronik wie zum Beispiel eines IMUs generiert wurde, ist, dass er hochfrequent Daten erhebt, hochdimensionale Daten sammelt und dadurch eine riesige Menge an Daten zusammenkommt. Einerseits ist genau das der gewünschte Aspekt, dass der Patient nicht mehr nur zu einem bestimmten Zeitpunkt in der Klinik, sondern öfter und kontinuierlicher untersucht wird. Andererseits macht das den Datensatz schwer zugänglich und erfordert ausgeklügelte Software, die die Rohdaten in verständliche und klinisch bedeutsame Information übersetzt (Kubota et al., 2016).

Das erfolgt mit Hilfe von selbstlernenden Algorithmen und künstlicher Intelligenz. Sie ist praktisch in der Lage in einem riesigen Datensatz eine Struktur zu erkennen, die von einem Menschen nicht zu erfassen wäre, und daraus Vorhersagen zu treffen, ob ein Patient zum Beispiel an einer Parkinson-Krankheit erkrankt ist oder nicht. Dabei programmiert sich der Vorhersage-Algorithmus allerdings anhand von zugeordneten Daten selbst. Das heißt, man muss die Software mit einem großen Datensatz füttern, bei dem vorher ein Arzt bestimmt hat, zu welcher Gruppe von Patienten (z.B. PK, RBD oder gesunde Kontrolle) ein Datenpunkt jeweils gehört, damit die Software dann daraus lernt. Die deduzierten Muster können dann auf andere Daten angewendet werden und so Patienten anhand ihrer Datensätze den verschiedenen Gruppen zugeordnet werden (Bishop, 2006).

Dabei ergeben sich aber eine Reihe von Problemen. Neben klassischen Einflussfaktoren auf die Messungen, wie Zeitpunkt, Tagesform oder Medikamenteneinnahme, ergeben sich auch spezifische Probleme beim Arbeiten mit „Machine-Learning-Algorithmen“. Exemplarisch beschreibt das Konzept des „Überfütterns“ (engl. „Overfitting“) und „Unterfütterns“ (engl. „Underfitting“) eines Modells die Schwierigkeit, die richtige Menge an Komplexität in einem „Machine-Learning-Modell“ zu erreichen, ohne Hintergrundrauschen oder andere Störfaktoren für klinisch relevante Information zu halten. Oder aber

ein Modell ist zu einfach und die Vorhersagen des Modells werden unpräzise. Außerdem ist die Plausibilität und Übertragbarkeit der Ergebnisse in großem Maße von der Datenqualität abhängig, die zu kontrollieren nicht einfach ist, wenn man beim „Home-Assessment“ einen Großteil der Kontrolle über den Versuchsaufbau abgibt (Kubota et al., 2016).

Diese Schwierigkeiten erfordern ein extrem kleinschrittiges Vorgehen in der Entwicklung und stetige Kontrolle auf Plausibilität und Validität. Insbesondere deshalb, weil bei der Verwendung von selbstlernenden Algorithmen nicht sofort ersichtlich ist, auf welche Parameter sich die Zuordnungsentscheidung stützt. Hierzu will die folgende Studie mit einer Subanalyse des Datensatzes einen Beitrag leisten.

3 Fragestellung

In der Vergangenheit konnte gezeigt werden, dass mit Hilfe von Smartphones, ihrer integrierten Sensortechnik und mit der entsprechenden Analyse durch ausgeklügelte statistische „Machine-Learning“-Programme, Parkinsonpatienten von Gesunden und Patienten mit REM-Schlaf-Verhaltensstörung mit hoher Treffsicherheit unterschieden werden konnten (Arora et al., 2018). In einer Kooperation mit dem Oxford Parkinson’s Disease Center haben wir den Versuchsaufbau an einer eigenen Kohorte an der Universität Marburg getestet, mit dem Ziel, die Ergebnisse einerseits zu reproduzieren und zu bestätigen, sowie andererseits eine eingehendere Analyse der erhobenen Daten durchzuführen, um Zusammenhänge und Relevanz einzelner Merkmale auf den Entscheidungsprozess insgesamt zu untersuchen. Davon erhofften wir uns ein tieferes Verständnis der Datenstruktur und Informationen darüber, welche Merkmale besonders sinnvoll zu messen sind, um in Zukunft noch zielgerichteter untersuchen zu können.

Zu diesem Zweck betrachteten wir die Testleistung (engl. „Test-Performance“) insgesamt, evaluierten, welche Merkmale den größten Einfluss auf die Entscheidungsfindung haben, und ob man zwischen diesen Merkmalen in den einzelnen Gruppen einen linearen Zusammenhang herstellen kann, im Sinne sich kontinuierlich verschlechternder Körperfunktionen.

4 Material und Methoden

Zur durchgeführten Studie liegt ein positives Votum der Ethikkommission des Fachbereichs Medizin der Philipps-Universität Marburg vom 27. Oktober 2015 mit dem Aktenzeichen 116/15 vor, welches bestätigt, dass das Studienprotokoll im Einklang mit nationaler Gesetzgebung und der Deklaration von Helsinki steht. Alle Teilnehmer willigten schriftlich nach ausführlicher mündlicher und schriftlicher Aufklärung in die Teilnahme an der Studie ein.

4.1 Finanzierung

Die Finanzierung der Studie erfolgte aus den Ausstattungsmitteln der Hertie-Senior-Forschungsprofessur für Neurowissenschaften, die seit 2013 die Forschungsarbeit von Herrn Prof. Dr. med. Dr. h.c. Wolfgang Oertel unterstützt. Eine finanzielle Aufwandsentschädigung der Probanden erfolgte nicht. Reisekosten und Kosten für Übernachtungen wurden ggf. übernommen.

4.2 Ein- und Ausschlusskriterien

Eingeschlossen wurden Patienten oder gesunde Personen im Alter zwischen 18 und 80 Jahren. Die männlichen oder weiblichen Probanden mussten in der Lage sein, ihr schriftliches Einverständnis zur freiwilligen Teilnahme an der Studie zu geben, und die Risiken und Komplikationen der Untersuchung zu verstehen und zu beurteilen. Diese Voraussetzungen galten für alle Kohorten. Ausgeschlossen aus der Studie wurden alle Personen, die nicht zur Einwilligung fähig waren.

4.2.1 Kohorte „RBD“

Probanden, die in die Kohorte der Patienten mit REM-Schlaf-Verhaltensstörung (RBD) eingeschlossen wurden, mussten die Diagnose einer isolierten REM-Schlaf-Verhaltensstörung gemäß den Kriterien der International Classification of Sleep Disorders erhalten haben (American Academy of Sleep Medicine, 2005).

4.2.2 Kohorte „PK“

Probanden, die in die Kohorte der Patienten mit Parkinson-Krankheit (PK) eingeschlossen wurden, mussten die Diagnose Parkinson-Krankheit anhand der Kriterien der Leitlinie der deutschen Gesellschaft für Neurologie erhalten haben (DGN, 2016). Ausgeschlossen wurden Patienten mit einem symptomatischen Parkinson-Syndrom (bsp. medikamentös-induziert, vaskulär bedingt im Rahmen einer subcortikalen arteriosklerotischen Enzephalopathie).

4.2.3 Kohorte „Gesunde Kontrollen“

Probanden, die in die Kohorte der gesunden Kontrollpersonen eingeschlossen wurden, durften weder die Diagnose REM-Schlaf-Verhaltensstörung noch die Diagnose Parkinson-Krankheit erhalten haben. Andere Erkrankungen galten nicht als Ausschlusskriterium. Sofern Diagnosen verfügbar waren beziehungsweise die Patienten Auskunft erteilten, wurden diese protokolliert.

4.2.4 Abbruchkriterien

Ein Abbruch der Studie war und ist für den Teilnehmer jederzeit möglich, ohne dass ihm dadurch Nachteile entstehen. Weiterhin ist der Abbruch der Studie aus gesundheitlichen, technischen oder anderen Gründen durch den Versuchsleiter möglich.

4.3 Applikation

Die Daten wurden mit der Applikation „HopkinsPD“ erhoben. Hierbei wurde darauf geachtet, dass die Applikation auf die neuste Version aktualisiert wurde. Der überwiegende Teil der Messungen wurde mit Version 1.7.9 durchgeführt. Die App besteht aus sieben Mini-Tests, die im Folgenden im Detail beschrieben werden sollen:

1. Stimmtest („Voice“):

Bei dem ersten Mini-Test zur Stimmcharakteristik handelt es sich um eine ca. 20 Sekunden dauernde Aufzeichnung der Stimme durch das integrierte Mikrofon im Telefon. Der Proband startet die Messung durch Antippen eines Buttons auf dem Touchscreen des Smartphones und führt das Smartphone anschließend – ähnlich wie beim Telefonieren – in die Nähe des Mundes. Daraufhin wird tief eingatmet

und so lange wie möglich „aaahh“ gesagt. Nach ca. 20 Sekunden wird die Aufzeichnung automatisch von der App beendet und der Startbildschirm des zweiten Tests erscheint automatisch auf dem Bildschirm des Handys.

2. Gleichgewichtstest („Balance“):

Bei dem zweiten Mini-Test wird der Patient aufgefordert, einen Ort aufzusuchen, an dem er frei stehen kann. Dort soll er, nachdem er die Aufzeichnung wieder mit einem Klick auf den Test-starten-Button beginnt, das Handy in der Hosentasche verstauen. Nach ca. 3 Sekunden vibriert das Handy und signalisiert damit, dass die Aufzeichnung startet. Der Proband muss nun nur im Raum stehen, ohne sich festzuhalten oder sich abzustützen. Nach ca. 20 weiteren Sekunden ist die Aufzeichnung beendet und das Smartphone signalisiert mit einer weiteren Vibration, dass die Messung abgeschlossen ist und springt wiederum in den Startbildschirm des nächsten Tests.

3. Gangtest („Gait“):

Der Test zur Gangcharakteristik läuft nach dem gleichen Prinzip wie der Gleichgewichtstest ab. Der Proband wird aufgefordert einen Ort aufzusuchen, an dem er möglichst ungehindert ca. 20 Meter geradeaus gehen kann. Ist das nicht über die volle Strecke möglich, sollen die Teilnehmenden nach der maximal möglichen Strecke einmal umdrehen und die Strecke wieder zurückgehen. Nach dem Antippen des Start-Buttons soll das Handy wiederum in der Hosentasche verstaut werden und nach einem Vibrationssignal soll der Proband nun die vorgegebene Strecke zurücklegen. Ist die Messung beendet, erfolgt ein weiteres Vibrationssignal und das Smartphone zeigt den nächsten Test an.

4. Geschicklichkeitstest („Finger Tapping“):

Der Proband wird gebeten das Smartphone auf eine gerade Oberfläche zu legen und startet wieder durch Antippen des Test-Starten-Buttons die Aufzeichnung. Auf dem Bildschirm erscheinen nun zwei Rechtecke A und B, die nun abwechselnd mit Zeige- und Mittelfinger der dominanten Hand in möglichst gleichmäßigem Rhythmus angetippt werden sollen. Ähnlich wie beim Klavier spielen soll, abhängig davon welche Hand verwendet wird, mit dem Zeigefinger immer Rechteck A und mit dem Mittelfinger immer Rechteck B angetippt werden oder andersherum. Das Smartphone zeichnet die vom Touchscreen erfassten Impulse auf und nach 20 Sekunden wird die Aufzeichnung von der App automatisch beendet.

5. Reaktionstest („Reaction Time“):

Beim Test zum Reaktionsvermögen erscheint nach dem Klick auf den Test-starten-Button in unregelmäßigen Abständen eine Schaltfläche auf der „Drück mich!“ steht. Diese soll so schnell wie möglich angetippt werden und der Finger soll, solange die Schaltfläche sichtbar ist, auf dieser belassen werden. Verschwindet die Schaltfläche wieder, soll auch der Finger wieder vom Touchscreen abgehoben werden. Erscheint die Schaltfläche erneut, wiederholt der Proband das Drücken

und Halten des Buttons. Das Smartphone erfasst über den Touchscreen die Reaktionszeiten der Probanden bei den unterschiedlich langen Präsentationen des Buttons.

6. Ruhetremor-Test („Rest Tremor“):

Beim sechsten Test zum Ruhetremor wird der Proband dazu aufgefordert, sich aufrecht hinzusetzen und das Handy in seiner tremordominanten Hand festzuhalten. Ist kein Tremor vorhanden, sollen die Patienten eine Hand ihrer Wahl benutzen. Nach Antippen des Start-Buttons signalisiert das Smartphone erneut mit einem Vibrationssignal den Start der Messung. Nun soll der Patient mit geschlossenen Augen von 100 rückwärts zählen. Nach ca. 45 Sekunden wird die Messung mit einem Vibrationssignal beendet und der nächste Test erscheint auf dem Bildschirm.

7. Haltetremor-Test („Postural Tremor“):

Der letzte Test des Messdurchlaufs testet den Tremor beim Vorhalten des Armes. Der Proband drückt den Test-Starten-Button und streckt in sitzender Position seinen Arm aus. Er hält das Handy in der gleichen Hand, mit der er auch den vorhergegangenen Test durchgeführt hat, nun ausgestreckt vor sich und mit einem Vibrationssignal beginnt die Messung. Nach ca. 45 Sekunden signalisiert die App mit einem weiteren Vibrationssignal, dass die Messung beendet ist. Der Proband sollte nach Möglichkeit den Arm die komplette Zeit der Messung ausgestreckt lassen. Sollte das nicht möglich sein, kann er die Messung auch abbrechen und den Arm sinken lassen. Nun verarbeitet das Smartphone die generierten Daten der vorangegangenen Messungen und es erscheint eine „Fertig-Schaltfläche“. Wird diese gedrückt, springt die App wieder in den Startbildschirm.

Im Startbildschirm der App kann man neben dem aktiven Testdurchlauf auch noch das passive Monitoring starten. Dieses sollen die Studienteilnehmer im Alltag während der Wachstunden aktivieren und dann das Handy am Körper tragen. Mit Antippen des Bewegungsaufzeichnung-Starten-Buttons wird die im Gerät verbaute IMU aktiviert. Mit Hilfe der so generierten Aufzeichnungen werden dann von der Applikation spezifische Bewegungsmuster erfasst und später ausgewertet. Das Smartphone zeigt in der App an, wann das passive Monitoring zuletzt durchgeführt wurde und ob das Monitoring im Moment aktiviert oder deaktiviert ist. Die Applikation erhebt keinerlei personenbezogene Daten wie Name oder Geburtsdatum. Die erhobenen Zielparameter werden verschlüsselt gespeichert und elektronisch auf einen projektbezogenen, gesicherten Server übermittelt. Informationen zur Entschlüsselung der pseudonymisierten Daten stehen nur projektbezogenen wissenschaftlichen Mitarbeitern zur Verfügung.

4.4 Technische Grundlagen der Testfunktionen

Das verwendete Testprotokoll entspricht weitestgehend dem aus der Universität Oxford bekannten *Procedere* (Arora et al., 2015). Dieses Testprotokoll, welches Stimm-, Gleichgewichts-, Gang-, Geschicklichkeits- und Reaktionstest enthielt, wurde um zwei Tests zur Tremoranalyse erweitert (Arora et al., 2018), wobei die im Smartphone integrierten IMUs zur Messung genutzt wurden (Barrantes et al., 2017). Diese Tests wurden von den Wissenschaftlern in Oxford (Arora et al., 2018) in der oben beschriebenen Art in die Applikation integriert, die Daten wurden verschlüsselt und mit einem Zeitstempel versehen auf eine sichere Online Datenbank hochgeladen.

4.5 Smartphones

Alle Messungen sind mit Android Smartphones durchgeführt worden.

4.5.1 HTC Desire 310

In einer ersten Beschaffungsmaßnahme wurden drei Smartphones vom Typ HTC Desire 310 beschafft, die auf der Android Version 4.2.2 laufen. Von diesen drei Handys wurde eines für alle klinischen Referenzmessungen verwendet.

4.5.2 Wiko Lenny 2

In einer zweiten Beschaffungsmaßnahme wurden sieben Smartphones vom Typ Wiko Lenny 2 beschafft, die auf der Android Version 5.1 laufen.

4.6 Probanden-Information

Jedem teilnehmenden Probanden wurde ein Informationsblatt mit allgemeinen Studieninformationen ausgehändigt, welches Informationen zur Durchführung, dem Ablauf und dem Sinn und Zweck der Studie enthielt. Weiterhin erhielt jeder Proband eine detaillierte Durchführungsanleitung, die Schritt für Schritt durch den Testablauf in der App führt. Grundkenntnisse in der Bedienung eines Smartphones wurden vorausgesetzt und falls nicht vorhanden, wurden die Probanden in der Bedienung bei der ersten Messung und dem Demo-Durchlauf angeleitet.

4.7 Vorgehen und Ablauf

4.7.1 Rekrutierung

Die Rekrutierung der Probanden für die Kohorten RBD und PK erfolgte aus den Spezial- und Studienambulanzen des Universitätsklinikums Marburg sowie von den neurologischen Stationen. Hierbei wurden im Rahmen der routinemäßigen klinischen Untersuchung durch einen Facharzt für Neurologie verschiedene klinische Daten erhoben, mit denen die Smartphone-Messungen in Relation gesetzt wurden. Diese klinischen Scores, insbesondere die UPDRS- („Unified Parkinson`s Disease Rating Scale“) und RBD-SQ („RBD-Screening-Questionnaire“)-Werte wurden, soweit vorhanden, nachträglich aus den Arztbriefen entnommen und protokolliert.

4.7.2 Einweisung

Nach umfassender Aufklärung und Einwilligung der Patienten in die Teilnahme an der Studie wurde den Probanden die Handhabung der Leih-Smartphones ausführlich erklärt. Daraufhin wurde mit den Probanden ein Testdurchlauf im Demo-Modus der App durchgeführt. Hierbei zeichnete die App keine Daten auf, sondern es wurde einmal die Benutzeroberfläche und der Ablauf der sieben verschiedenen Tests, sowie des passiven Monitorings demonstriert und erklärt. Eine ausführliche Aufklärung über den Ablauf der Studie, sowie eine detaillierte Anleitung der Durchführung der Tests und Bedienung der Applikation, wurde den Probanden ausgehändigt (s. Anhang).

4.7.3 Klinische Messung

Direkt im Anschluss an die ausführliche Einweisung und den Demo-Durchlauf wurde eine klinische Referenzmessung unter Aufsicht durchgeführt, um beurteilen zu können, ob und wie der Proband mit der Steuerung der App und der Durchführung der Tests zurechtkommt. Diese erste Referenzmessung diente auch dazu, im Vergleich mit den Messungen zu Hause, abschätzen zu können, ob den Probanden bei den Messungen in Eigenregie grobe Fehler unterlaufen sind.

4.7.4 Heimmessung

Die Heimmessung erfolgte nach ausführlicher Einweisung in der Klinik bei den Probanden zu Hause. Die Probanden machten eine Woche lang jeden Tag vier Mal einen Durchlauf durch alle sieben Tests. Die Messungen sollten über den Tag verteilt stattfinden, um ein möglichst genaues Abbild des Verlaufs der Tagesform zu erhalten. Nach Möglichkeit und Zeit der Patienten sollten die Messungen am Tag nach der ambulanten Vorstellung oder dem stationären Aufenthalt in der Klinik morgens beginnen und dann eine Woche andauern. Je nach Möglichkeit und Terminen der Probanden wurde aber auch ein späterer Beginn gestattet. Auf Wunsch wurden in das Studien-Smartphone Wecker oder Erinnerungen programmiert, die eine regelmäßige Messung unterstützen sollten.

Parallel zur Messung mit den aktiven Tests wurde in den Wachstunden auch das passive Monitoring so oft wie möglich durchgeführt. Die Studienteilnehmer entschieden zu jeder Zeit selbst, wie oft und wie lange sie das passive Monitoring aktivierten.

4.7.5 Rücksendung und Upload der generierten Daten

Nach Abschluss der Messungen wurde das Studien-Smartphone, inklusive Ladegerät, in einem mitgegebenen, adressierten und frankierten Briefumschlag an die Klinik zurückgesandt. Dort wurde das Smartphone mit dem W-LAN verbunden und die aufgezeichneten Daten auf die Studienserver hochgeladen. Wenn die Probanden das Smartphone zu Hause mit ihrem W-LAN verbunden, wurden die generierten Daten automatisch auf die Studienserver hochgeladen.

4.8 Auswertung der Daten

Die Auswertung von Daten, die mit herkömmlichen Smartphones unter wenig kontrollierten Bedingungen bei Probanden zu Hause erfolgt, bringt einige Unsicherheiten mit sich. Um Artefakt reiche Datenabschnitte so gut es geht auszuschließen, wurde ein automatisierter Teilungsalgorithmus benutzt, der Artefakte sozusagen ausschneidet. Hierbei wurden Aufzeichnungen nur berücksichtigt, wenn alle sieben Tests in ausreichender zeitlicher Nähe zueinander durchgeführt wurden (Arora et al., 2018).

4.9 Statistische Auswertung der Ergebnisse

Ziel war es festzustellen, ob die objektive Erfassung der motorischen Parameter erlaubte, die untersuchte Person einer der drei Diagnose-Gruppen (gesunde Kontrollen, RBD, PK) zuzuordnen, sowie die Art und Weise, wie die Zuordnung durch den Algorithmus erfolgt, zu erforschen.

Drei verschiedene Vergleiche wurden angestrebt und die dazugehörige Entscheidungssicherheit betrachtet. Die Kontrollgruppe und RBD-Kohorte, die RBD-Kohorte und Parkinson-Kohorte, sowie die Parkinson-Kohorte und die Kontrollgruppe wurden mittels zwei verschiedenen statistischen Testverfahren, nämlich der 10-fachen Kreuzvalidierung (10-fach-KV) und der „Leave-One-(Recording)-Out“ (LORO) Kreuzvalidierung, verglichen. Die Zuordnungsentscheidung des Algorithmus wurde mit den Ergebnissen der ärztlichen Untersuchung verglichen, welche als wahre Ergebnisse angesehen wurden. Ziel der Auswertung war es, die motorischen Symptome objektiv quantifizieren zu können, die Merkmale mit dem größten Beitrag zur Entscheidungsfindung des Algorithmus zu identifizieren und die Genauigkeit der Entscheidung zu analysieren.

Um genauere Aussagen über den Algorithmus, nämlich seine Art zu arbeiten und eine Zuordnungsentscheidung zu treffen, machen zu können, wurde für jedes Test-Set (RBD versus Kontrollen LORO und 10-fach-KV, PK versus Kontrollen LORO und 10-fach-KV, RBD versus PK LORO und 10-fach-KV) eine Rangfolge erstellt, welche Merkmale am meisten zur Entscheidungsfindung beitragen. Dieser Aspekt erfolgt üblicherweise mittels Algorithmen, die redundante und unnütze Informationen entfernen (Arora et al., 2018).

Um die drei Probanden-Gruppen voneinander zu unterscheiden, wurden, wie schon bei Arora et al. beschrieben, „Random Forests“ benutzt (Arora et al., 2018). Bei der Klassifizierung mittels „Random Forests“ entstehen beim Trainieren des Algorithmus hunderte verschiedene voneinander unabhängige Entscheidungs-Bäume (vgl. Abb. 1), die alle eine Zuordnungsentscheidung zu den jeweiligen Gruppen treffen. Die Gruppe, die häufiger gewählt wird, bestimmt das Resultat (Breiman, 2001).

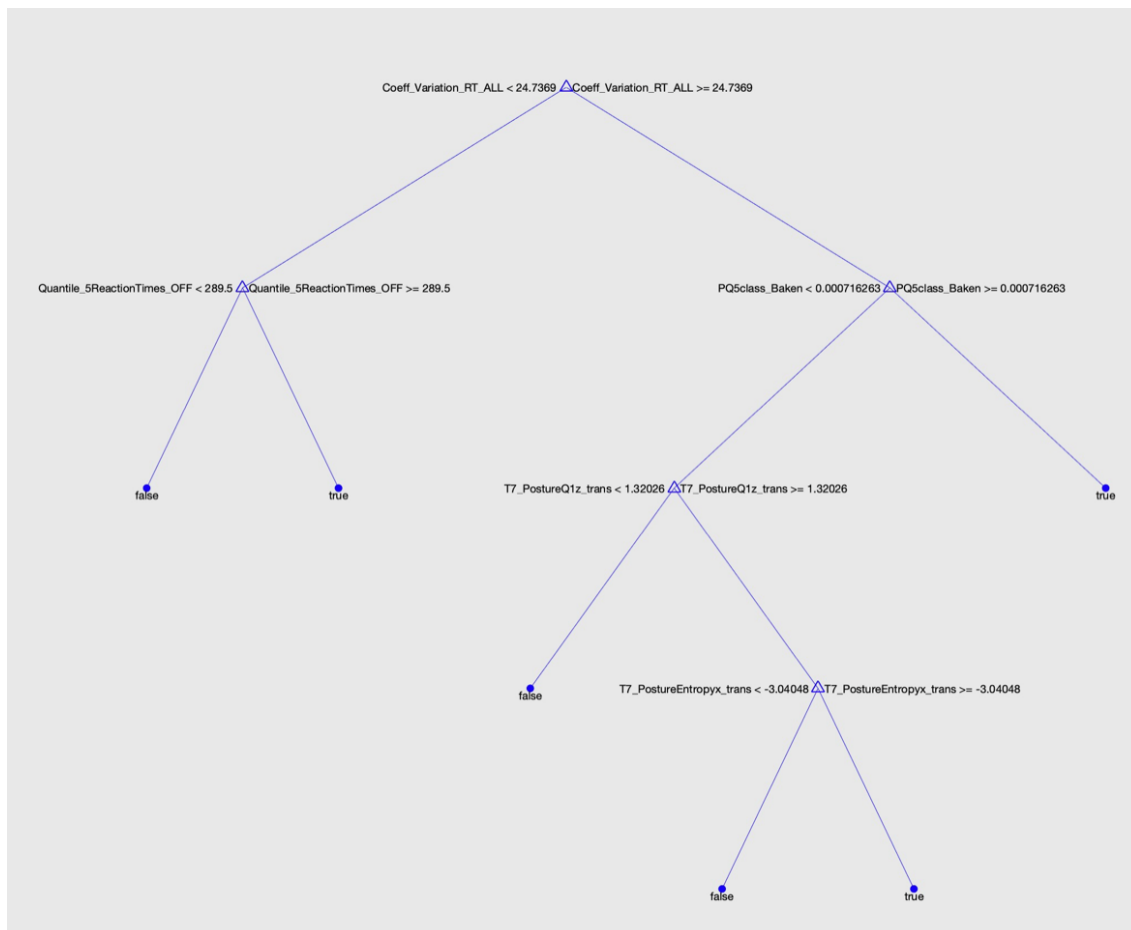


Abbildung 1: „Random Forests“

Exemplarische Darstellung eines Ausschnitts des „Random Forests“ zur Zuordnung der Teilnehmer zu den jeweiligen Gruppen.

Abbildung zur Verfügung gestellt vom Oxford Parkinson's Disease Center

Die interne Validierung erfolgte an unserem Datensatz mit zwei verschiedenen Methoden der Kreuzvalidierung (KV), der 10-fachen-KV und „Leave-One-Recording-Out“-KV.

Dabei werden die Messungen als unabhängig voneinander betrachtet und der Datensatz wird wiederholt in eine Trainingsgruppe und eine Testgruppe geteilt. Bei der 10-fachen-KV wird der Datensatz so geteilt, dass 90% der zufällig ausgewählten Messungen zum Training benutzt werden und 10% als Testmenge.

Bei der LORO-KV wird eine einzelne Aufzeichnung als Testmenge und der gesamte Rest zum Training genutzt.

Auf dieser Grundlage wurde analysiert, ab wie vielen inkorporierten Merkmalen ein Optimum in der Testleistung erreicht wird. Das heißt also, ab wann noch mehr Information keine bessere Leistung unseres Algorithmus verursachte.

Um nun beurteilen zu können, ob für die drei Vergleiche dieselben Merkmale eine hohe Wichtigkeit (engl. „Importance“) haben oder ob jeweils andere Merkmale in unterschiedlichen Krankheitsstadien aussagekräftiger sind, erstellten wir eine gruppenübergreifende

Rangfolge der Merkmale. Dabei bildeten wir aus den Rängen der Wichtigkeits-Rangfolge der einzelnen Merkmale für die Vergleiche RBD vs. gesunde Kontrollen, PK vs. gesunde Kontrollen und PK vs. RBD die Summe aus den Rängen dieser drei paarweisen Vergleiche. Je niedriger die Summe, desto höher ist nun die Wichtigkeit über alle Vergleiche hinweg. Mit diesem Procedere untersuchten wir über alle drei Test-Sets hinweg die Wichtigkeit der einzelnen Merkmale aus den beiden Versuchsreihen LORO und 10-fach-KV. Wir betrachteten im Gegensatz zu vorangegangenen Studien die so identifizierten Merkmale gesondert und verglichen sie über alle drei Probandengruppen hinweg miteinander. Dabei erhofften wir uns erklärende Informationen darüber, welche Symptome oder Unter Aspekte von Symptomen besonders geeignet sind, um sie mittels Smartphone-Sensoren zu untersuchen, und wie es in Zukunft gelingen könnte, die Komplexität in der Analyse zu reduzieren.

Die statistische Analyse der Algorithmus-Leistung und die Rangfolge der Merkmale wurde für uns von der Arbeitsgruppe um Siddharth Arora am Oxford Parkinson's Disease Center mit Matlab (Version 2016b) erstellt. (Arora et al., 2018)

Die weitergehende Analyse zur gruppenübergreifenden Wichtigkeit und das Erstellen der Grafiken erfolgten mit dem Statistik Programm R[®] und Microsoft Excel.

5 Ergebnisse

5.1 Demographische und klinische Charakteristik der Probanden

	RBD N=43	PK N=19	gesunde Kon- trollen N=34
<i>Alter (Jahre)</i> MW±SD	64,9 ± 6,8	65,3 ± 8,0	64,1 ± 6,5
<i>Alter (Jahre)</i> Median	65,1	67,7	63,1
<i>weiblich</i> n (%)	4 (9,3)	0 (0,0)	21 (61,8)
<i>männlich</i> n (%)	39 (90,7)	19 (100)	13 (38,2)

Tabelle 1: Demographische Daten

Tabelle 1 zeigt Anzahl und Alter der Probanden, sowie die Verteilung der Geschlechter in den Gruppen zum Zeitpunkt der Messung.

Abkürzungen: RBD = REM-Schlaf-Verhaltensstörung, PK = Parkinson-Krankheit, MW = Mittelwert, SD = Standardabweichung

	RBD N=43	PK N=19	gesunde Kontrol- len N=34
<i>MoCA</i>	27,4 ± 1,8	28,8 ± 1,2	29,2 ± 0,6
<i>RBD-SQ</i>	9,3 ± 2,7	6,3 ± 3,8	2,6 ± 1,9
<i>PDNMS- Skala</i>	7,8 ± 4,0	10,6 ± 4,8	5,4 ± 4,3
<i>BDI</i>	7,9 ± 8,6	13 ± 12,8	11 ± 5,8
<i>UPDRS Teil III</i>	1,8 ± 1,8	17,1 ± 11,9	3,1 ± 5,8

Tabelle 2: Klinische Daten (Mittelwerte ± Standardabweichung)

Die Tabelle zeigt die Mittelwerte und Standardabweichung der angegebenen Parameter.

Abkürzungen: RBD = REM-Schlaf-Verhaltensstörung, PK = Parkinson-Krankheit, MoCA = Montreal Cogni-

tive Assessment, RBD-SQ = REM-Sleep Behavior Disorder Screening Questionnaire, PDNMS-Skala = Parkinson's Disease Non-Motor Symptoms-Scale, BDI = Beck Depressions Inventar, UPDRS Teil III = Unified Parkinson's Disease Rating Scale Teil III.

Tabellen 1 und 2 zeigen die demographischen und klinischen Daten aller eingeschlossenen Personen. In der untersuchten Kohorte waren Patienten mit Parkinson-Krankheit im mittel älter als Patienten mit RBD und Kontrollen. Insgesamt waren Patienten mit RBD und Parkinson-Krankheit auch häufiger männlich als die gesunden Kontrollen. Aus Tabelle 2 sind klinische Unterschiede zwischen den Gruppen im Hinblick auf psychische und motorische sowie vegetative Einschränkungen erkennbar. Eine schwerere kognitive Einschränkung zeigt sich bei keiner der Kohorten. Patienten mit Parkinson-Krankheit zeigen ein schlechteres Abschneiden im UPDRS Teil III, also ein schlechteres Abschneiden in der motorischen Untersuchung. Ebenso war die im Mittel erreichte Punktzahl im RBD-SQ in der RBD-Kohorte mit 9,3 am höchsten, gefolgt von der PK-Kohorte mit 6,3 und den gesunden Kontrollpersonen mit 2,6. Tabelle 3 zeigt die Anzahl der tatsächlich ausgewerteten Messungen bestehend aus allen sieben Tests. Im Mittel trug jeder RBD-Patient 19,9 Aufnahmen zu unserem Datensatz bei. Parkinsonpatienten trugen im Mittel 16,3 Aufnahmen bei und liegen damit noch unter dem Wert für die Kontroll-Kohorte, in der im Mittel 17,6 Aufnahmen pro Patienten durchgeführt wurden. Außerdem ist aus Tabelle 3 ersichtlich, dass einige Patienten ausgeschlossen werden mussten. Dies geschah vor allem auf Grund von zu wenigen übermittelten Daten. Insgesamt entstanden so 547 Aufnahmen aus der Kohorte der Kontrollen, 757 Aufnahmen aus der Kohorte RBD und 309 Aufnahmen aus der Kohorte PK. Jede dieser Aufnahmen besteht aus den 7 Tests, aus denen wiederum 998 Merkmale extrahiert wurden, die den Rohdatensatz bilden, welcher mittels der Kreuzvalidierungsverfahren untersucht wurde.

	<i>Teilnehmer N</i>	<i>Aufnahmen N</i>	<i>Aufnahmen / Teilnehmer</i>
<i>Kontrollen</i>	31	547	17,6
<i>RBD</i>	38	757	19,9
<i>PK</i>	19	309	16,3

Tabelle 3: ausgewertete Daten

Die Tabelle zeigt die Anzahl (N) der eingeschlossenen und ausgewerteten Teilnehmer und wie viele Testdurchläufe diese absolviert haben, sowie die durchschnittliche Menge an Testdurchläufen pro Teilnehmer.

Abkürzungen: RBD = REM-Schlaf-Verhaltensstörung, PK = Parkinson Krankheit

5.2 Testleistung

Die Testleistung, im Sinne der Sicherheit der Zuordnung zu den drei Kohorten Kontrollen, RBD und PK nach Training und Test mittels LORO-KV und 10-fach-KV, wurde nun mittels ROC-Kurven (engl.: „receiver operating characteristic“) und den zugehörigen AUCs (engl.: „area under the curve“) beurteilt und graphisch dargestellt. Es wurden je nach Menge der analysierten Merkmale mehrere ROC-Kurven erstellt, die die statistische Sicherheit der Entscheidung für die geeignete Wahl der Testparameter darstellen.

5.2.1 Testleistung nach LORO-Kreuzvalidierung und 10-fach-Kreuzvalidierung für die Unterscheidung zwischen RBD und Kontrollen

Die Testleistung bei Verwendung der 30 wichtigsten Merkmale ist, wie in der Vorarbeit von Arora et. al. (Arora et al., 2018) beschrieben, vergleichbar mit der Verwendung aller 998 Merkmale. Die Abbildungen 2 und 3 zeigen die Testleistung für die Unterscheidung zwischen RBD-Probanden und gesunden Kontrollpersonen bei Verwendung der Top 10, Top 20, Top 30, Top 40 und Top 50 Merkmale. Abbildung 2 zeigt ROC-Kurven und die korrespondierenden AUCs bei einer Validierung mittels LORO-KV.

Eine ROC-Kurve nahe der Diagonalen deutet auf einen Zufallsprozess hin. Je besser die Testleistung ist, in unserem Fall je sicherer die Zuordnung zu den getesteten Kohorten, desto steiler ist der Anstieg der ROC-Kurve und desto größer ist auch die zugehörige AUC. Die AUC kann dementsprechend Werte zwischen 0,5 und 1 annehmen, wobei 0,5 einer zufälligen Zuordnung entspricht und 1 eine hundertprozentige Zuordnungssicherheit ausdrückt.

Hierbei wird deutlich, dass sich die Testleistung zu Beginn verbessert, je mehr Merkmale in die Analyse miteinbezogen wurden. Die AUC wuchs von 0,69 bei 10 Merkmalen auf 0,73 bei 30 und 40 betrachteten Merkmalen, wo sie ihr Optimum erreichte. Danach ließ die Testleistung wieder nach und die AUC sank wieder auf 0,71 bei 50 betrachteten Merkmalen.

Testleistung des Modells nach Training/Validierung mittels LORO Kreuzvalidierung an den Kohorten RBD und Kontrollen

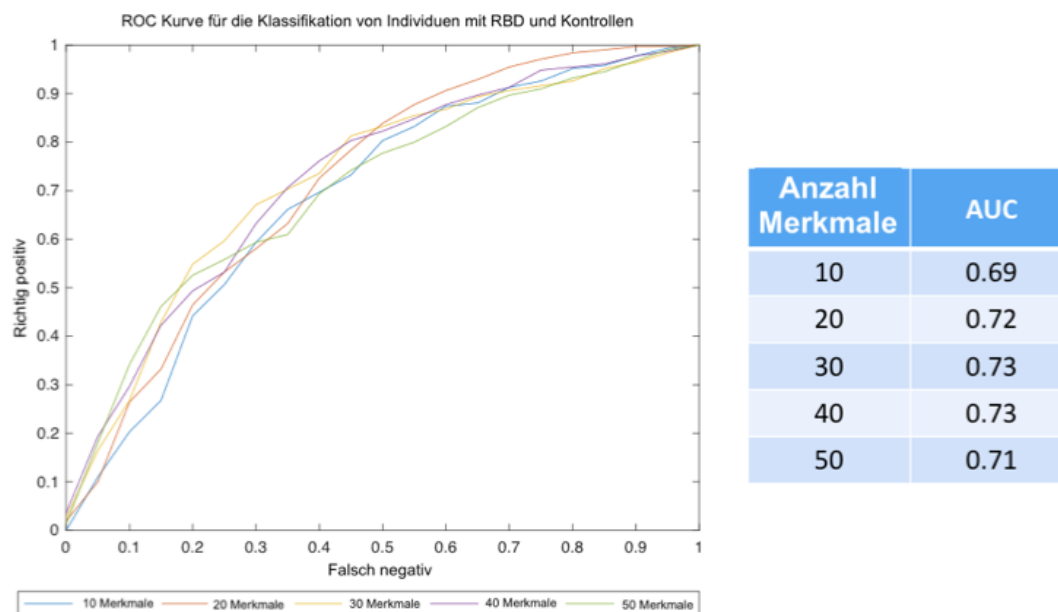


Abbildung 2: Testleistung LORO RBD vs. Kontrollen

Die Abbildung zeigt eine ROC-Kurve für eine Zuordnung zur RBD-Kohorte oder der Kontroll-Kohorte nach Trainieren eines Modells mittels LORO. Die Kurven zeigen die Leistung des Modells bei Zuordnungsentscheidungen anhand der 10, 20, 30, 40 bzw. 50 wichtigsten Merkmale.

Abkürzungen: AUC = Area under the Curve, ROC = Receiver Operating Characteristic, LORO = Leave One Recording Out, RBD = REM-Schlaf-Verhaltensstörung

Abbildung erstellt vom Oxford Parkinson's Disease Center, Datenerhebung und Bearbeitung von Moritz Beilharz

Betrachtet man nun Abbildung 3, die ROC-Kurven und die korrespondierenden AUCs bei einer Validierung mittels 10-fach-KV darstellt, erkennt man eine deutlich bessere Testleistung als bei LORO-Kreuzvalidierung, also eine sicherere Zuordnung zu den beiden Kohorten RBD oder gesunde Kontrollen. Auch hier verbessert sich die Testleistung, je mehr Merkmale in die Analyse eingeschlossen wurden, wobei eine gute Sicherheit bei ca. 30 inkorporierten Merkmalen erreicht wurde. Die höchste Zuordnungssicherheit und eine AUC von 0,96 wurde bei 50 inkorporierten Merkmalen erreicht, die niedrigste bei 10 Merkmalen mit einer AUC von 0,85. Alle AUCs der 10-fach-KV waren größer als die der LORO-KV, was also einer deutlich besseren Testleistung entspricht.

Testleistung des Modells nach Training/Validierung mittels 10-fach Kreuzvalidierung an den Kohorten RBD und Kontrollen

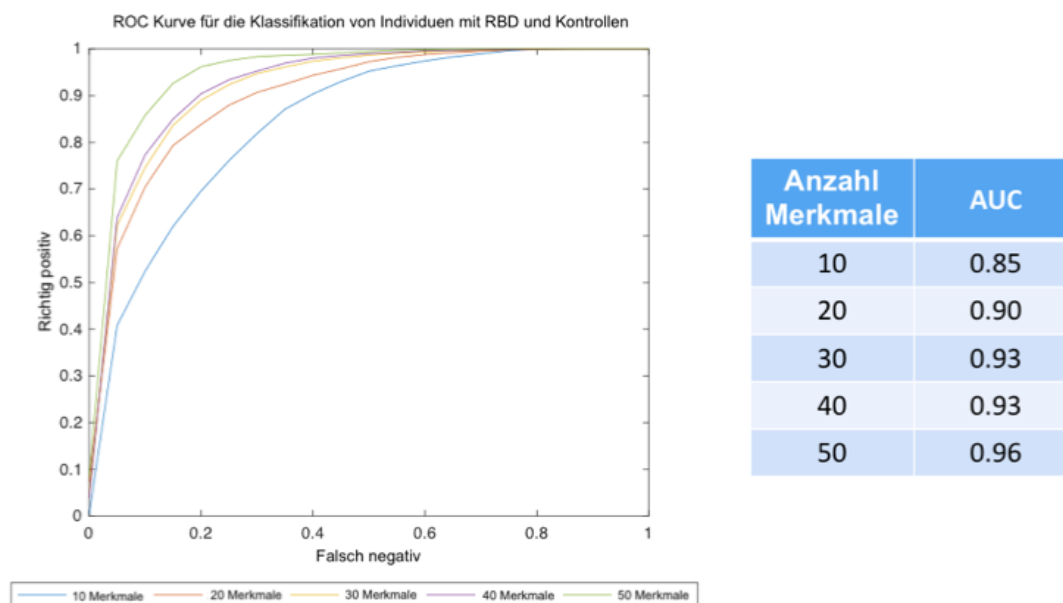


Abbildung 3: Testleistung 10-fach-Kreuzvalidierung RBD vs. Kontrollen

Die Abbildung zeigt eine ROC-Kurve für eine Zuordnung zur RBD-Kohorte oder der Kontroll-Kohorte nach Trainieren eines Modells mittels 10-fach-KV. Die Kurven zeigen die Leistung des Modells bei Zuordnungsentscheidungen anhand der 10, 20, 30, 40 bzw. 50 wichtigsten Merkmale.

Abkürzungen: AUC = Area under the Curve, ROC = Receiver Operating Characteristic, 10-fach-KV = 10-fache Kreuzvalidierung. RBD = REM-Schlaf-Verhaltensstörung.

Abbildung erstellt vom Oxford Parkinson's Disease Center, Datenerhebung und Bearbeitung von Moritz Beilharz

Insgesamt zeigt sich eine sehr gute Zuordnungssicherheit im Vergleich zur LORO-KV.

5.2.2 Herkunft der einflussreichsten Merkmale RBD vs. Kontrollen

Abbildung 4 zeigt korrespondierend dazu, aus welchem der sieben Tests der App die untersuchten wertvollsten Merkmale stammen. Exemplarisch bedeutet das für den Vergleich RBD versus Kontrollen, dass bei der 10-fach-KV 13 der Top 30 Merkmale, also 43%, auf den Bereich des Stimmtests entfallen. Damit hat die Stimmanalyse in diesem Test-Set die meisten Merkmale beigetragen, gefolgt von der Reaktionszeit. Bei der LORO-KV hat auch der Gleichgewichtstest einen wichtigen Stellenwert, wohingegen sich bei der 10-fach-KV die Testentscheidung noch mehr auf Haltetremor-Merkmale stützt. Bei der 10-fach-KV zeigt sich, dass die wichtigsten Merkmale aus nur insgesamt drei Tests stammen, wohingegen bei der LORO-KV insgesamt Merkmale aus vier Tests unter den Top 30 Merkmalen vertreten sind.

Herkunft der Top 30 Unterscheidungsmerkmale für RBD vs. Kontrollen

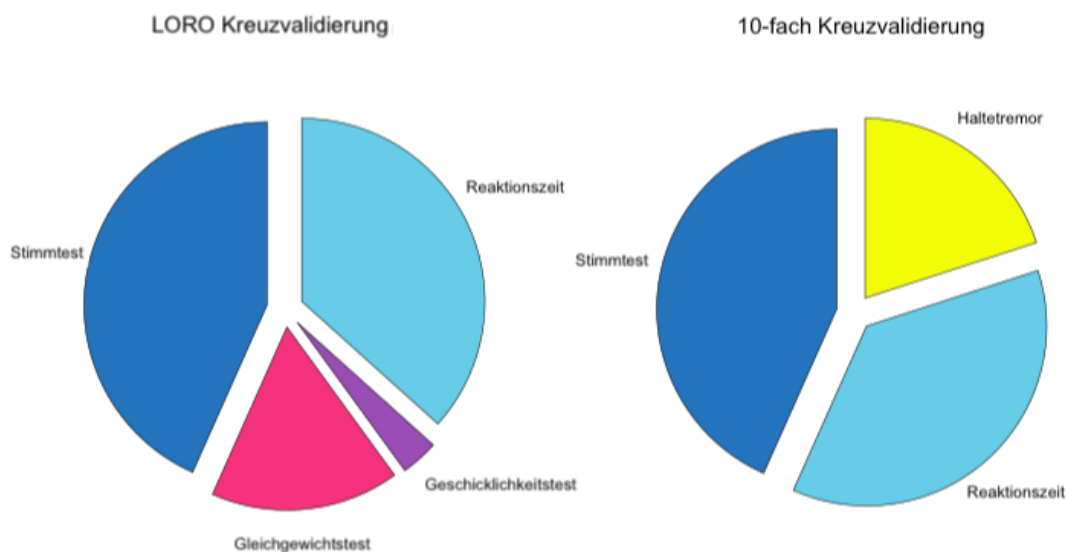


Abbildung 4: Unterscheidungsmerkmale RBD vs. Kontrollen

Abbildung 4 zeigt die Herkunft der 30 wichtigsten Merkmale für die Zuordnungsentscheidung RBD vs. Kontrollen aus den verschiedenen Tests der Applikation. Hieraus ergeben sich Hinweise für Wertigkeit und Einfluss der verschiedenen Tests.

Abkürzungen: RBD = REM-Schlaf-Verhaltensstörung, LORO = Leave One Recording Out.

Abbildung erstellt vom Oxford Parkinson's Disease Center, Datenerhebung und Bearbeitung von Moritz Beilharz

5.2.3 Testleistung nach LORO-Kreuzvalidierung und 10-fach-Kreuzvalidierung für die Unterscheidung zwischen PK und Kontrollen

Äquivalent zum Vergleich zwischen RBD und Kontrollen zeigen die nachfolgenden Grafiken 5-7 die Analyse für die Test-Sets zwischen der PK- und der Kontroll-Kohorte. Auch hier ist eine deutlich sicherere Zuordnung für die 10-fach-KV zu erkennen. Bei der LORO-Kreuzvalidierung (Abbildung 5) wurden AUCs zwischen 0,68 bei 10 und 50 Merkmalen bis 0,76 bei 30 Merkmalen erreicht.

Testleistung des Modells nach Training/Validierung mittels LORO Kreuzvalidierung an den Kohorten PK und Kontrollen

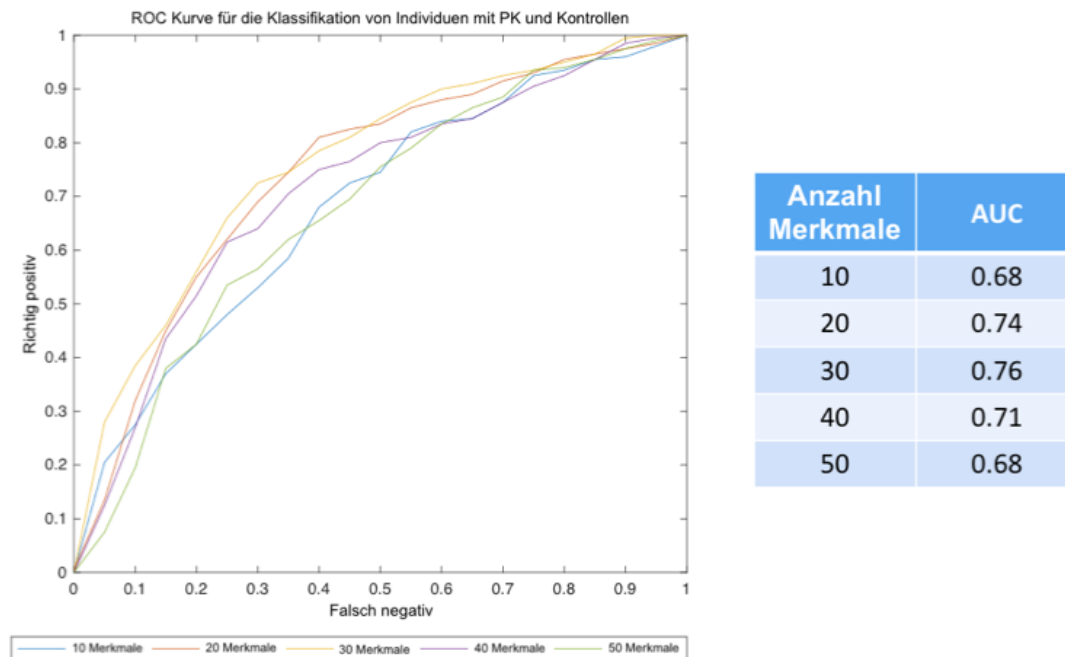


Abbildung 5: Testleistung LORO PK vs. Kontrollen

Die Abbildung zeigt eine ROC-Kurve für eine Zuordnung zur PK-Kohorte oder der Kontroll-Kohorte nach Trainieren eines Modells mittels LORO. Die Kurven zeigen die Leistung des Modells bei Zuordnungsentscheidungen anhand der 10, 20, 30, 40 bzw. 50 wichtigsten Merkmale.

Abkürzungen: AUC = Area under the Curve, ROC = Receiver Operating Characteristic, LORO = Leave One Recording Out, PK = Parkinson-Krankheit.

Abbildung erstellt vom Oxford Parkinson's Disease Center, Datenerhebung und Bearbeitung von Moritz Beilharz

Bei der 10-fach-KV (Abbildung 6) wurden AUCs zwischen 0,82 bis 0,92 erreicht, wobei sich die Analyse hierbei verbesserte, je mehr Merkmale eingeschlossen wurden.

Testleistung des Modells nach Training/Validierung mittels 10-fach Kreuzvalidierung an den Kohorten PK und Kontrollen

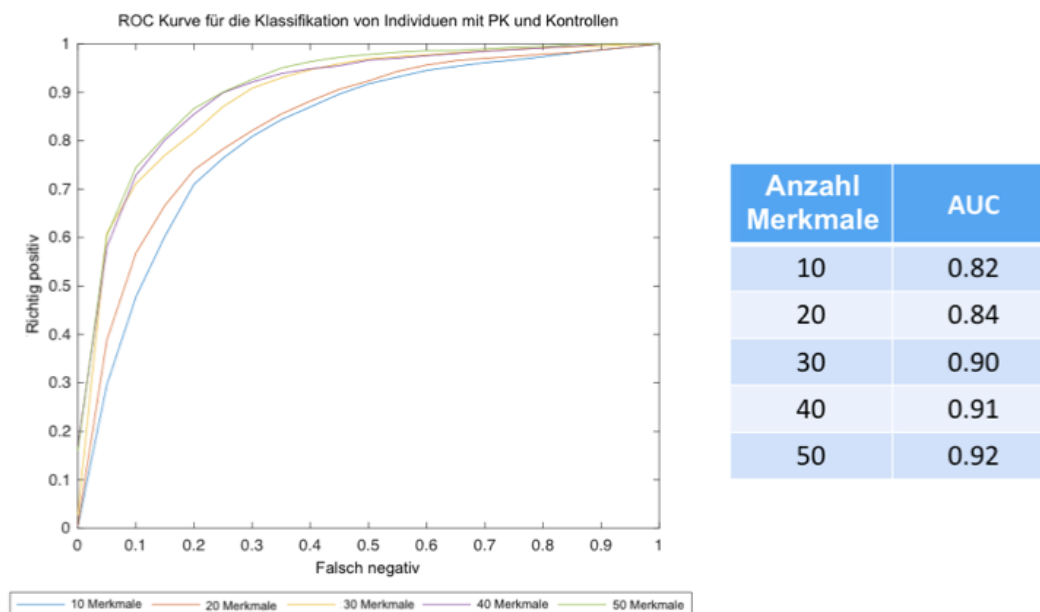


Abbildung 6: Testleistung 10-fach-Kreuzvalidierung PK vs. Kontrollen

Die Abbildung zeigt eine ROC-Kurve für eine Zuordnung zur PK-Kohorte oder der Kontroll-Kohorte nach Trainieren eines Modells mittels 10-fach-KV. Die Kurven zeigen die Leistung des Modells bei Zuordnungsentscheidungen anhand der 10, 20, 30, 40 bzw. 50 wichtigsten Merkmale.

Abkürzungen: AUC = Area under the Curve, ROC = Receiver Operating Characteristic, 10-fach-KV = 10-fache Kreuzvalidierung, PK = Parkinson-Krankheit.

Abbildung erstellt vom Oxford Parkinson's Disease Center, Datenerhebung und Bearbeitung von Moritz Beilharz

Es zeigt sich erneut nur bei der 10-fach-KV eine akzeptable Zuordnungssicherheit, wobei die Testleistung insgesamt etwas schlechter ist als bei der Unterscheidung zwischen RBD und Kontrollen. Dies ist sicherlich unter anderem der in der PK-Kohorte geringeren Teilnehmerzahl bzw. der daraus resultierenden geringeren Zahl an Aufnahmen geschuldet.

5.2.4 Herkunft der einflussreichsten Merkmale PK vs. Kontrollen

Vergleicht man die Zusammensetzung der Top 30 Merkmale in Abbildung 7 mit Abbildung 4, so fällt auf, dass sich die Gruppen, aus denen die Merkmale mit dem höchsten Einfluss stammen, von denen der Analyse RBD vs. Kontrollen unterscheiden.

Beide Auswertungen LORO und 10-fach-KV gewichten die Analyse von Haltetremor Symptomen stärker als beispielsweise die Stimmanalyse. Die Reaktionszeit spielt erneut eine wichtige Rolle, ebenso wie die Analyse des Haltetremors. Es zeigt sich, dass diesmal die Top 30 Merkmale der 10-fach-KV etwas diverser zusammengesetzt sind.

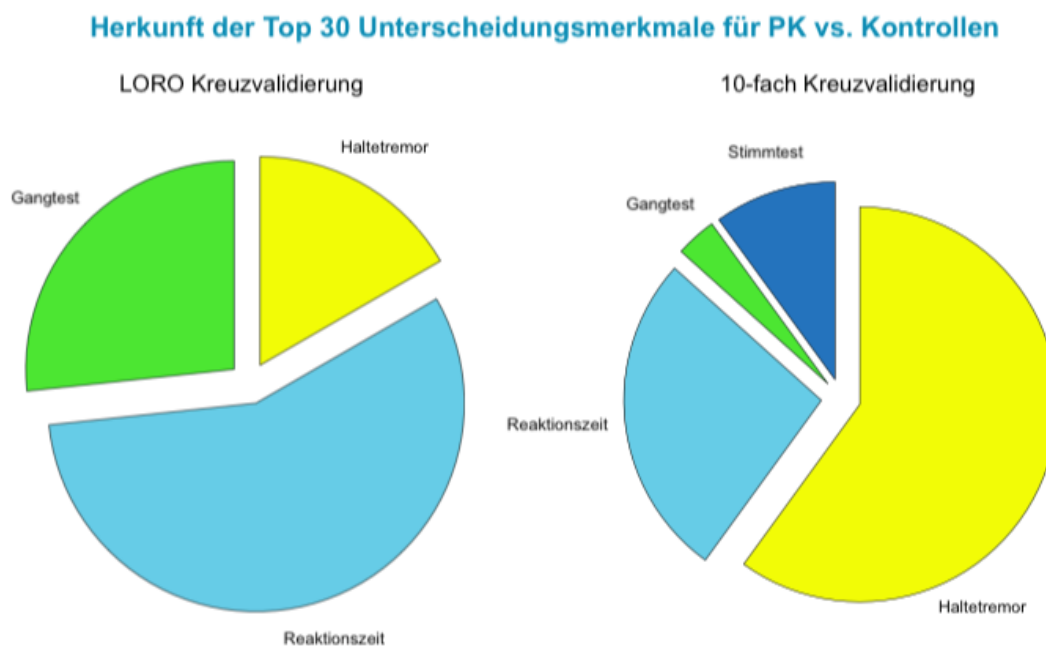


Abbildung 7: Separationsmerkmale PK vs. Kontrollen

Abbildung 7 zeigt die Herkunft der 30 wichtigsten Merkmale für die Zuordnungsentscheidung PK vs. Kontrollen aus den verschiedenen Tests der Applikation. Hieraus ergeben sich Hinweise für Wertigkeit und Einfluss der verschiedenen Tests.

Abkürzungen: PK = Parkinson-Krankheit, LORO = Leave One Recording Out.

Abbildung erstellt vom Oxford Parkinson's Disease Center, Datenerhebung und Bearbeitung von Moritz Beilharz

5.2.5 Testleistung nach LORO-Kreuzvalidierung und 10-fach-Kreuzvalidierung für die Unterscheidung RBD vs. PK

Die Abbildungen 8-10 zeigen die Ergebnisse des Vergleichs der beiden Kohorten RBD und PK. Äquivalent zu den vorangegangenen Tests hat die LORO-Kreuzvalidierung eine relativ schlechte Zuordnungssicherheit gezeigt, wohingegen die 10-fach-KV nach ca. 30 Merkmalen wieder sehr zuverlässig wurde und eine $AUC > 0,9$ erreichte. AUCs der LORO-KV rangierten zwischen 0,63 und 0,71, während AUCs der 10-fach-KV kontinuierlich von 0,83 zu 0,95 zunahmen, je mehr Merkmale in die Analyse eingeschlossen wurden.

Testleistung des Modells nach Training/Validierung mittels LORO Kreuzvalidierung an den Kohorten RBD und PK

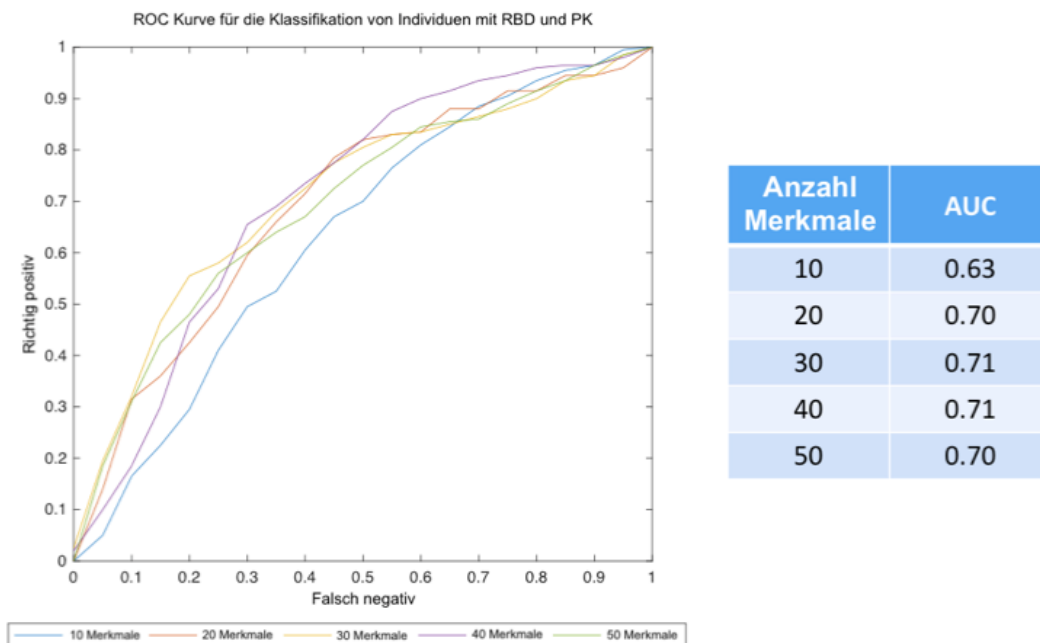


Abbildung 8: Testleistung LORO PK vs. RBD

Die Abbildung zeigt eine ROC-Kurve für eine Zuordnung zur PK-Kohorte oder der RBD-Kohorte nach Trainieren eines Modells mittels LORO. Die Kurven zeigen die Leistung des Modells bei Zuordnungsentscheidungen anhand der 10, 20, 30, 40 bzw. 50 wichtigsten Merkmale.

Abkürzungen: AUC = Area under the Curve, ROC = Receiver Operating Characteristic, LORO = Leave One Recording Out, PK = Parkinson-Krankheit, RBD = REM-Schlaf-Verhaltensstörung.

Abbildung erstellt vom Oxford Parkinson's Disease Center, Datenerhebung und Bearbeitung von Moritz Beilharz

Testleistung des Modells nach Training/Validierung mittels 10-fach Kreuzvalidierung an den Kohorten RBD und PK

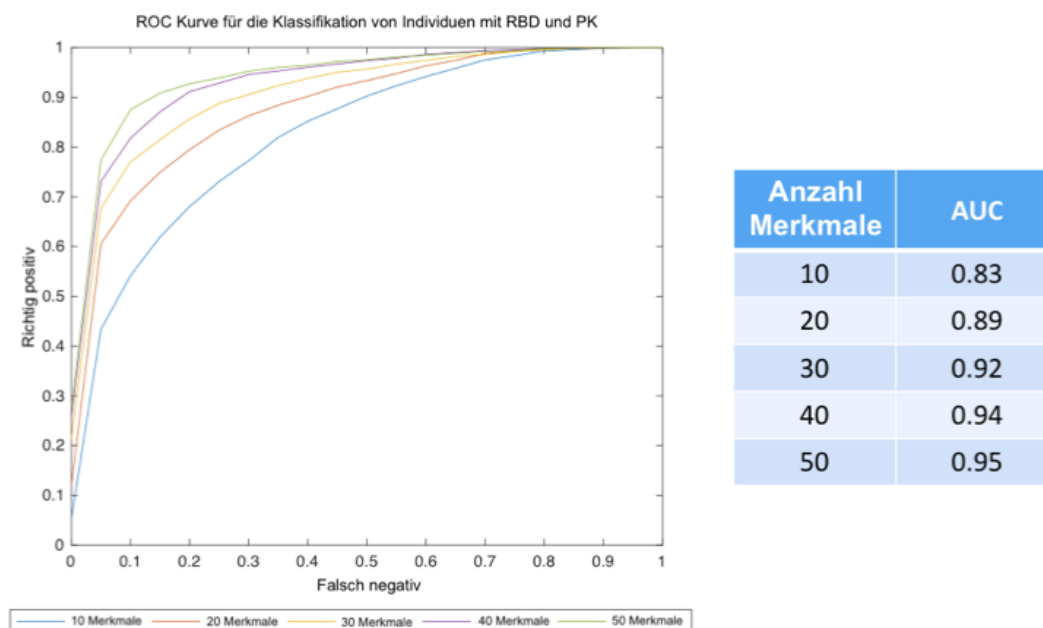


Abbildung 9: Testleistung 10-fach-Kreuzvalidierung PK vs. RBD

Die Abbildung zeigt eine ROC-Kurve für eine Zuordnung zur RBD-Kohorte oder der PK-Kohorte nach Trainieren eines Modells mittels 10-fach-KV. Die Kurven zeigen die Leistung des Modells bei Zuordnungsentscheidungen anhand der 10, 20, 30, 40 bzw. 50 wichtigsten Merkmale.

Abkürzungen: AUC = Area under the Curve, ROC = Receiver Operating Characteristic, 10-fach-KV = 10-fache Kreuzvalidierung, RBD = REM-Schlaf-Verhaltensstörung, PK = Parkinson-Krankheit.

Abbildung erstellt vom Oxford Parkinson's Disease Center, Datenerhebung und Bearbeitung von Moritz Beilharz

5.2.6 Herkunft der einflussreichsten Merkmale PK vs. RBD

Wie zu erwarten, setzen sich die Merkmale mit dem größten Einfluss auf die Entscheidungsfindung aus den einflussreichsten Merkmalen der vorangegangenen Testreihen zusammen. Sowohl die Stimmanalyse als auch die Tremoranalyse steuern einen großen Teil der wichtigen Merkmale bei. Die Reaktionszeitmessung verliert hingegen deutlich an Einfluss und ist nicht mehr unter den Top 30 Merkmalen mit dem höchsten Einfluss auf die Zuordnungsentscheidung zu finden. Die Top 30 Merkmale stammen bei der 10-fach-KV erneut aus nur drei Tests, wohingegen bei der LORO-KV fünf Tests unter den wichtigsten 30 Merkmalen vertreten sind.

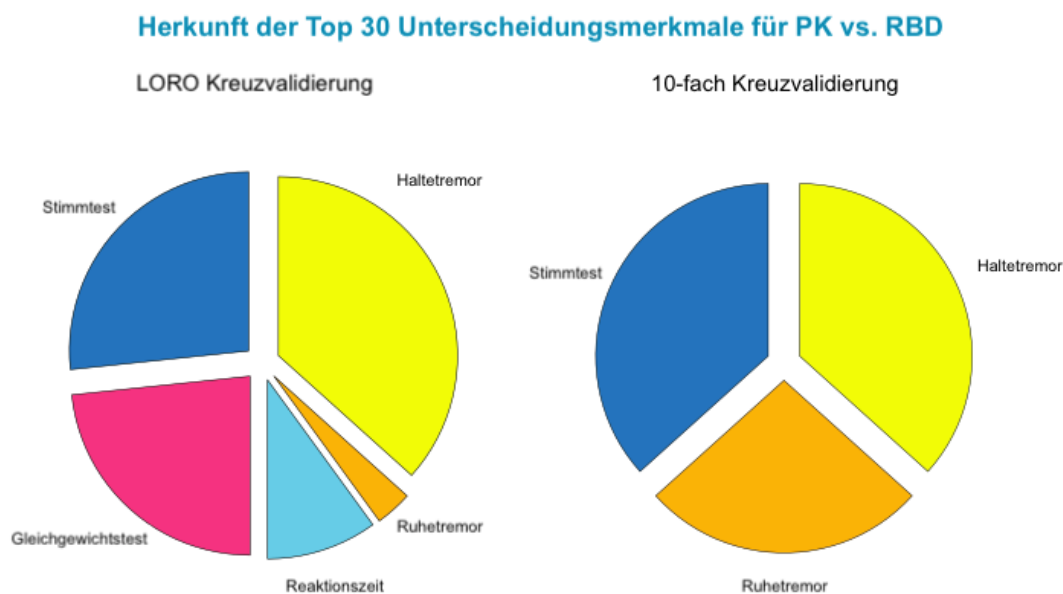


Abbildung 10: Separationsmerkmale PK vs. RBD

Abbildung 10 zeigt die Herkunft der 30 wichtigsten Merkmale für die Zuordnungsentscheidung PK vs. RBD aus den verschiedenen Tests der Applikation. Hieraus ergeben sich Hinweise für Wertigkeit und Einfluss der verschiedenen Tests.

Abkürzungen: PK = Parkinson-Krankheit, RBD = REM-Schlaf-Verhaltensstörung, LORO = Leave One Recording Out.

Abbildung erstellt vom Oxford Parkinson's Disease Center, Datenerhebung und Bearbeitung von Moritz Beilharz

Die Kreisdiagramme aus den Abbildungen 4, 7 und 10 geben zwar Aufschluss darüber, wie viele der 30 wichtigsten Merkmale einem bestimmten Mini-Test bzw. Analyse Cluster zuzuordnen sind, allerdings erlauben sie nur indirekt Rückschlüsse auf den tatsächlichen Einfluss, den die einzelnen Merkmale haben. Erstens ist es möglich, dass die einzelnen Merkmale redundante Informationen enthalten, die die Zuordnungssicherheit zwar marginal verbessern, aber dazu führen, dass der Einfluss eines Tests überschätzt wird. Zweitens ist es nicht möglich zu sagen, welcher Anteil der Entscheidungsfindung tatsächlich auf ein einzelnes Merkmal entfällt.

5.3 Aussagekraft und Wichtigkeit der Merkmale

Um hier mehr Klarheit zu schaffen, erstellten wir eine Wichtigkeits-Rangfolge der 998 Merkmale für jedes 10-fach-KV Modell. Außerdem bildeten wir die Summe der Ränge der Merkmale über alle drei Modelle hinweg, um zu schätzen, inwieweit ein einzelnes Merkmal Aussagen zu allen Gruppen erlaubt oder ob es nur zur Diskriminierung einer einzelnen Gruppe geeignet ist. Tabelle 4 zeigt die 30 wichtigsten Merkmale von allen 998 aus den Vergleichen RBD vs. gesunde Kontrollen, PK vs. gesunde Kontrollen und RBD

vs. PK, sortiert nach absteigender Wichtigkeit. Summiert man nun die Ränge der Merkmale und ordnet sie mit der niedrigsten Summe beginnend aufsteigend, erhält man eine Rangfolge, welche die Wichtigkeit der Merkmale über alle drei Vergleiche hinweg zeigt (Spalte „Merkmals-Rangliste gesamt“ in Tabelle 4). Das Merkmal „T7_PostureQ1x_trans“ hat in Tabelle 4 in Spalte 1 den Rang 27, in Spalte 2 den Rang 26 und in Spalte 3 den Rang 12, was in der Summe 65 ergibt. Damit hat es die niedrigste Summe der Ränge aller Merkmale gefolgt von „T7_PostureMeanx_trans“ mit der Summe 77. Daraus ergibt sich die höchste Aussagekraft zu allen drei Gruppen RBD, PK und Kontrollen.

	<i>RBD vs. Kontrollen</i>	<i>PK vs. Kontrollen</i>	<i>RBD vs. PK</i>	<i>Merkmals-Rangliste gesamt</i>
1	mean_FO	T7_PostureMeanSquaredEnergyx	RPDE	T7_PostureQ1x_trans
2	RPDE	T7_PostureMeanSquaredEnergyx_trans	mean_FO	T7_PostureMeanx_trans
3	max_FO	T7_PostureRootMeanSquarex	det_LT_TKEO_std_8_coef	T7_PostureMedianx_trans
4	median_FO	T7_PostureRootMeanSquarex_trans	det_LT_entropy_log_9_coef	T7_PostureQ3x_trans
5	app_det_TKEO_mean_7_coef	Diff_OnOff_RT	max_FO	T7_PostureQ1x
6	muFO	T7_PostureAR1x_trans	median_FO	T7_PostureMeanx
7	app_entropy_log_8_coef	MedianReactionTimes_OFF	app_det_TKEO_mean_7_coef	T7_PostureQ3x
8	Quantile_15ReactionTimes_OFF	Quantile_55ReactionTimes_OFF	det_entropy_log_9_coef	T7_PostureMedianx
9	Quantile_5ReactionTimes_ALL	Quantile_45ReactionTimes_OFF	app_entropy_log_8_coef	Diff_OnOff_RT
10	Quantile_5ReactionTimes_OFF	T7_PostureAR1x	GQ1	T7_PostureModex
11	Min_RT_OFF	T7_PostureInterQuartileRangex	T6_RestureQ3xyz	T7_PostureModex_trans
12	Quantile_25ReactionTimes_OFF	Quantile_65ReactionTimes_OFF	T7_PostureQ1x_trans	T7_PostureMeanTKEOx
13	det_entropy_log_9_coef	T7_PostureMeanx	T7_PostureModex_trans	T3_GaitureMeanTKEOy
14	Min_RT_ALL	T7_PostureQ3x_trans	T7_PostureMedianx_trans	T7_PostureQ3z
15	Quantile_15ReactionTimes_ALL	T7_PostureMedianx	T7_PostureMeanx	T3_GaitureQ3xyz
16	T7_PostureMeanth_CO	Quantile_25ReactionTimes_ALL	T7_PostureQ3x	T7_PostureMeanTKEOx_trans
17	Mode_RT_OFF	T7_PostureQ3x	T7_PostureMedianx	det_LT_entropy_log_9_coef
18	Quantile_35ReactionTimes_OFF	T7_PostureMeanx_trans	T7_PostureQ1x	T7_PostureQ3y
19	T7_PostureMeanth_CO_trans	T7_PostureModex	T7_PostureModex	T7_PostureMeanSquaredEnergyx_trans
20	det_LT_TKEO_std_8_coef	medMFCC9	T7_PostureMeanx_trans	T7_PostureMeanz
21	Mu_ExG_RT_OFF	T7_PostureMedianx_trans	T6_RestureMedianxyz	T7_PostureQ1y
22	app_entropy_log_9_coef	T7_PostureInterQuartileRangex_trans	T7_PostureQ3x_trans	T7_PostureMeany
23	det_LT_entropy_log_2_coef	T7_PostureQ1x	T6_RestureMeanxyz	T7_PostureRootMeanSquarex_trans
24	RelBandPower3	MeanReactionTimes_OFF	T6_RestureMeanSquaredEnergyxyz	app_det_TKEO_mean_7_coef
25	Quantile_45ReactionTimes_OFF	app_entropy_log_5_coef	T6_RestureMeanr_CO	T7_PostureMeanSquaredEnergyx
26	RelBandPower1	T7_PostureQ1x_trans	T6_RestureRootMeanSquarexyz	T3_GaitureMedianxyz
27	T7_PostureQ1x_trans	T7_PostureEntropyx_trans	muFO	T7_PostureMediany
28	T7_PostureQ1y_trans	RMS_RT_OFF	T6_RestureQ1y_trans	Min_RT_OFF
29	T7_PostureMediany_trans	T3_GaitureMeanr_CO_trans	T6_RestureModey_trans	Quantile_5ReactionTimes_OFF
30	T7_PostureMeany_trans	det_LT_entropy_log_6_coef	T7_PostureInterQuartileRangex	T7_PostureRootMeanSquarex

Tabelle 4: Top 30 Merkmale der 10-fachen Kreuzvalidierung

Tabelle 4 zeigt die Namen der 30 Merkmale mit dem höchsten Einfluss auf die Zuordnungsentscheidung in absteigender Reihenfolge für alle drei paarweisen Zuordnungen und insgesamt. Die im Text erwähnten Merkmale sind zur besseren Verständlichkeit hervorgehoben.

Abkürzungen: PK = Parkinson-Krankheit, RBD = REM-Schlaf-Verhaltensstörung.

Erkennbar ist hier schon, dass Test Nummer sieben (erkennbar am Präfix T7) zum Hal-tetremor (engl. „postural tremor“) einen großen Teil der wichtigsten Merkmale stellt, ins-besondere bei Betrachtung aller Gruppen in Spalte 4 (Merkmals-Rangliste gesamt). Dar-aus lässt sich eine verhältnismäßig hohe Aussagekraft zu allen Kohorten und Krankheits-bildern folgern. In den drei einzelnen Vergleichen zeigt sich zum Teil eine noch stärkere

Gewichtung von anderen Tests und Merkmalen. Hinweise auf die Gründe hierfür gibt die explorative Datenanalyse im Folgenden.

5.3.1 Aussagekraft der wichtigsten Merkmale RBD vs. Kontrollen

Wir führten eine explorative Datenanalyse durch, um mehr über die Arbeitsweise des „Machine-Learning“-Algorithmus zu erfahren. Die Merkmale mit der höchsten Wichtigkeit in der Rangfolge bieten sich dafür besonders an. Die explorative Datenanalyse erfolgte mittels Streudiagrammen und Boxplots, um zu illustrieren, wie die tatsächliche Struktur der Daten ist und ob sich visuell Tendenzen in der Datenstruktur ableiten lassen. Wir beschränkten uns dabei auf die Ergebnisse der 10-fach-KV, da, wie oben gezeigt, die Genauigkeit des Modells deutlich besser ist als bei der LORO-KV.

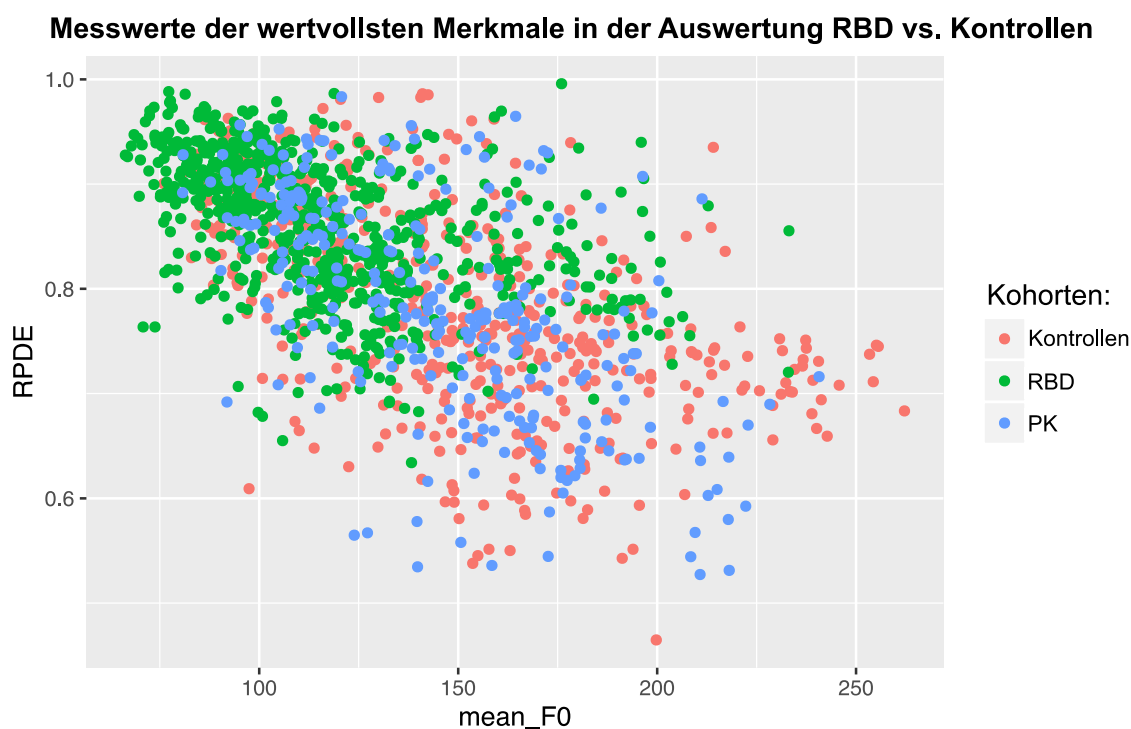


Abbildung 11: Streudiagramm der zwei wertvollsten Merkmale in der Auswertung RBD vs. Kontrollen

Abbildung 11 zeigt ein Streudiagramm, das die zwei Merkmale „RPDE“ und „mean_F0“ (beides Merkmale der Stimmanalyse) gegeneinander aufträgt. Diese hatten für den Vergleich RBD vs. Kontrollen die höchste Wichtigkeit, waren also am besten in der Lage zwischen RBD und Kontrollen zu unterscheiden. Jeder Punkt stellt einen Messwert dieser beiden Merkmale dar. Es werden die Messwerte aller drei Kohorten dargestellt, um mögliche Unterschiede in der Werteverteilung zu visualisieren. Es zeigt sich eine Ähnliche Werteverteilung für PK und Kontrollen, wohingegen sich der grüne Cluster für RBD vom Rest der Datenpunkte abhebt. Abkürzungen: PK = Parkinson-Krankheit, RBD = REM-Schlaf-Verhaltensstörung, RPDE = Recurrence Period Density Entropy.

Abbildung 11 zeigt ein Streudiagramm, in dem alle Datenpunkte, also jeder einzelne Testdurchlauf von jedem Probanden aller drei Kohorten, der beiden wichtigsten Merkmale

des Vergleichs zwischen RBD und Kontrollen gegeneinander aufgetragen sind. Die beiden wichtigsten Merkmale in diesem Fall sind „RPDE“ und „mean_F0“, die Werte aus der Stimmanalyse darstellen.

Hierbei bezeichnet „mean_F0“ in der Stimmanalyse die mittlere Grundfrequenz (Lindstrom, Ohlsson, Sjöholm, & Waye, 2010). „RPDE“ (engl. Recurrence Period Density Entropy) ist ein Maß für die Periodizität bzw. Dynamik eines Signals. Die „RPDE“ wird in der Stimmanalyse zur Abgrenzung der Umgebungsgeräusche und beispielsweise zur Analyse der Heiserkeit eingesetzt (Little, McSharry, Roberts, Costello, & Moroz, 2007).

Es zeichnen sich im Prinzip zwei Cluster ab, die man mit bloßem Auge voneinander differenzieren kann. In Rot sind die Datenpunkte für die Kontrollgruppe dargestellt, blau sind die Datenpunkte für die PK-Kohorte und in grün sieht man die Werte für die RBD-Kohorte. Auffällig ist, dass die Werte der RBD-Erkrankten mit einer höheren Wahrscheinlichkeit im linken oberen Quadranten liegen, die Werte für Kontrollen und PK aber ein sehr ähnliches Verteilungsmuster aufweisen. Dieser Befund bestätigt sich auch, wenn man Abbildung 12 betrachtet. Die Boxplots stellen die Verteilungen der Werte für die gleichen Parameter dar, wobei wiederum die Werte für die RBD-Kohorte von denen für PK und Kontrollen deutlich zu unterscheiden sind, nicht aber PK und Kontrollen untereinander.

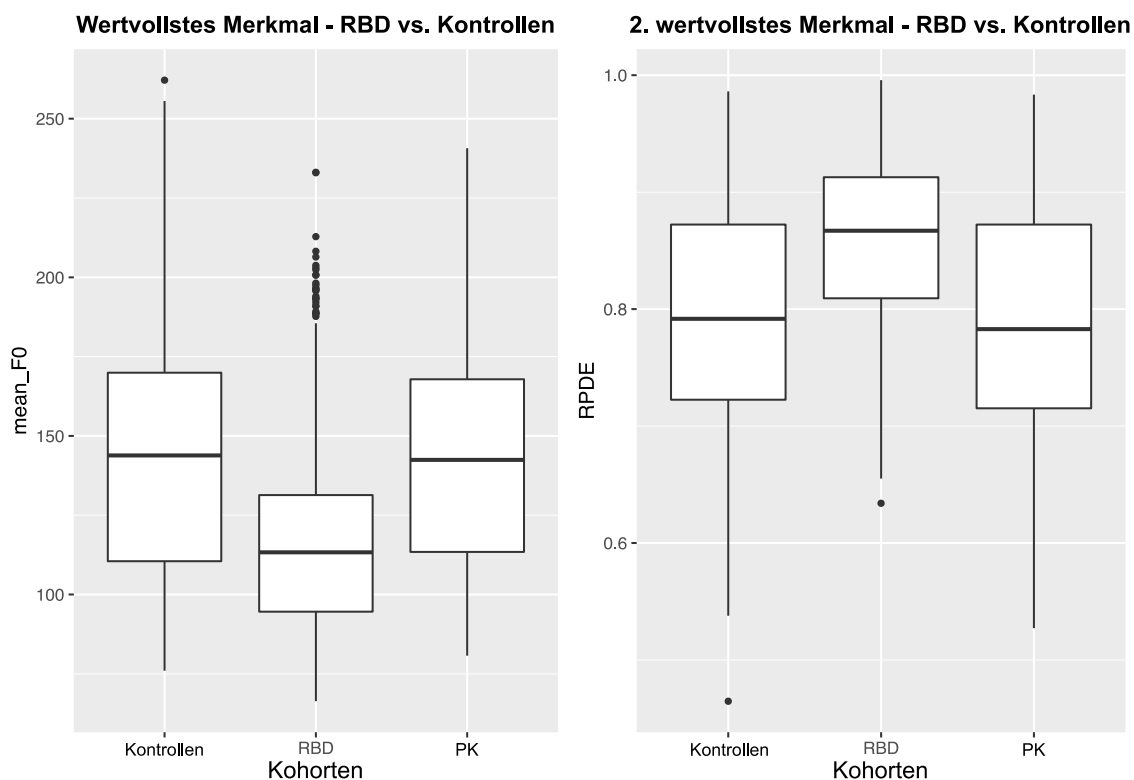


Abbildung 12: Boxplots wertvollste Merkmale RBD vs. Kontrollen

Abbildung 12 zeigt Boxplots, welche die Verteilung der Messwerte um die jeweiligen Mittelwerte der beiden Merkmale „mean_F0“ und „RPDE“ (beides Merkmale der Stimmanalyse) in den drei untersuchten Kohorten darstellen. Es zeigen sich deutliche Unterschiede in der Wertausprägung zwischen der RBD-Kohorte und den zwei anderen Kohorten.

Abkürzungen: RBD = REM-Schlaf-Verhaltensstörung, PK = Parkinson-Krankheit, RPDE = Recurrence Period Density Entropy.

Hieraus ergibt sich, dass mit diesen zwei Parametern „mean_F0“ und „RPDE“ die RBD Kohorte gut von den anderen beiden Kohorten Kontrollen und PK abgegrenzt werden kann. Eine Aussage kann also vor allem über die RBD Kohorte getroffen werden.

5.3.2 Aussagekraft zweier wichtiger Merkmale PK vs. Kontrollen

Abbildungen 13 und 14 zeigen zwei wichtige Merkmale, die mit dieser Analyse sinnvoll zu unterscheidenden sind, für den Vergleich der Kohorte PK mit der Kohorte Kontrollen. Teilweise werden in der Rangfolge der Merkmale redundante Informationen noch höher gewertet, da sie scheinbar die Leistung des Algorithmus verbessern. Allerdings lassen sich diese nicht sinnvoll visualisieren. Deshalb sind in den beiden Abbildungen das wichtigste Merkmal der Reaktionszeitmessung („Diff_OnOff_RT“) und ein wichtiges Merkmal der Messung des posturalen Tremors (Haltetremors) („T7_PostureRootMeanSquare“) gegeneinander aufgetragen. Im Streudiagramm stellt sich die Analyse schwierig dar. Auf den ersten Blick lassen sich keine spezifischen Cluster voneinander trennen, welche die hohe Wichtigkeit rechtfertigen würden.

Messwerte der wertvollen Merkmale in der Auswertung PK vs. Kontrollen



Abbildung 13: Streudiagramm der zwei wertvollen Merkmale in der Auswertung PK vs. Kontrollen
Abbildung 13 zeigt ein Streudiagramm, das die stark gewichteten Merkmale „Diff_OnOff_RT“ (ein Merkmal aus der Reaktionszeitmessung) und „T7_PostureRootMeanSquarex“ (ein Merkmal aus der Messung des posturalen Tremors) im Vergleich der Gruppen PK vs. Kontrollen gegeneinander aufträgt. Jeder Punkt stellt einen Messwert dieser beiden Merkmale dar. Es werden die Messwerte aller drei Kohorten dargestellt, um mögliche Unterschiede in der Werteverteilung zu visualisieren.

Abkürzungen: PK = Parkinson-Krankheit, RBD = REM-Schlaf-Verhaltensstörung.

Erst in den Boxplots (Abbildung 14) zeigt sich, dass es deutliche Unterschiede zwischen den beiden Kohorten gibt. Die Werte für „T7_PostureRootMeanSquarex“ und „Diff_OnOff_RT“ veränderten sich kontinuierlich von Kontrollen über RBD nach PK. Sie scheinen somit geeignet, um Aussagen über alle drei Kohorten zu treffen. Dies deckt sich auch mit der Wichtigkeits-Rangfolge für alle drei Vergleiche in Tabelle 4 Spalte 4 (Merkmals-Rangliste gesamt). Dort sind beide Merkmale in den Top 30 der Gesamtauswertung gelistet. Durch die scheinbar kontinuierliche Veränderung der Werte von Kontrollen über RBD zu PK scheinen diese beiden Merkmale insbesondere geeignet, eine Krankheitsprogression über alle Kohorten hinweg zu beobachten.

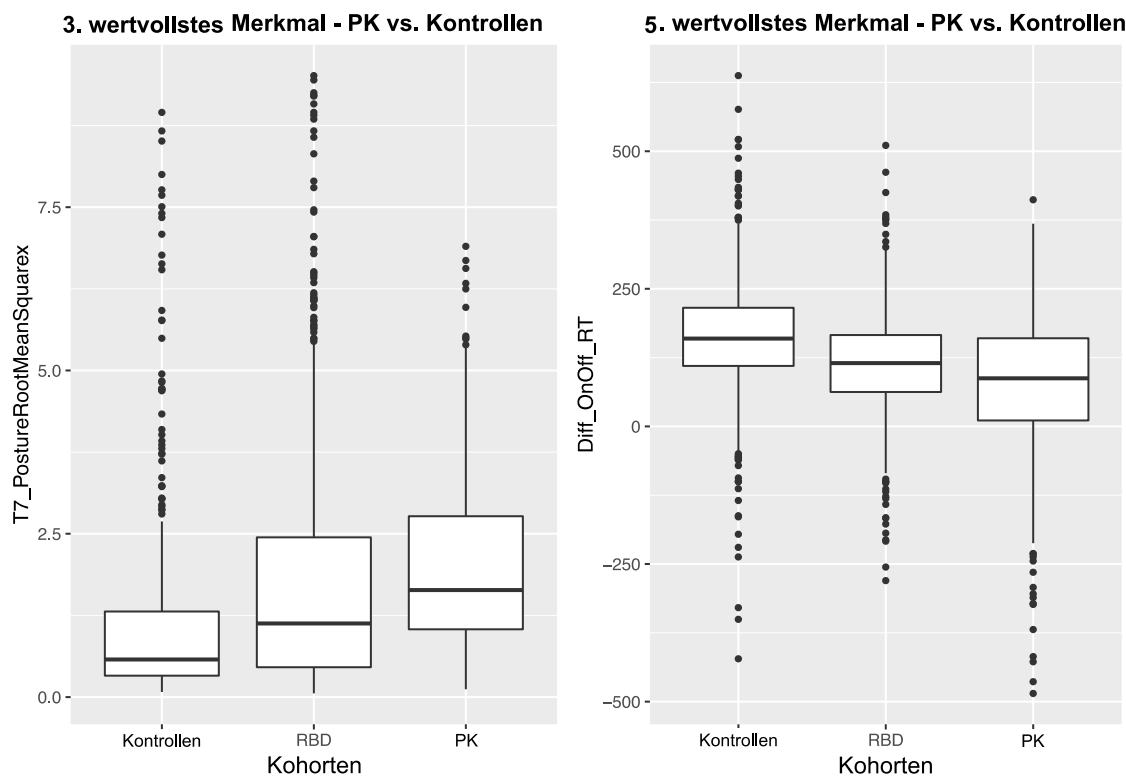


Abbildung 14: Boxplots wertvolle Merkmale PK vs. Kontrollen

Abbildung 14 zeigt Boxplots, welche die Verteilung der Werte um die jeweiligen Mittelwerte der beiden Merkmale „T7_PostureRootMeanSquare“ (ein Merkmal aus der Messung des posturalen Tremors) und „Diff_OnOff_RT“ (ein Merkmal aus der Reaktionszeitmessung) in den jeweiligen Gruppen darstellen. Es findet sich eine deutliche Tendenz von niedrigen Werten in der Kohorte der gesunden Kontrollen, hin zu höheren Werten bei der RBD-Kohorte bis zur höchsten Wertausprägung in der PK-Kohorte für „T7_PostureRootMeanSquare“. Umgekehrt zeigen sich bei „Diff_OnOff_RT“ bei den Kontrollen tendenziell höhere Werte als bei RBD bis hin zu den niedrigsten Werten bei PK. Dies kann im Sinne einer kontinuierlichen Verschlechterung der Messwerte mit Fortschreiten der Erkrankung interpretiert werden.

Abkürzungen: RBD = REM-Schlaf-Verhaltensstörung, PK = Parkinson-Krankheit.

5.3.3 Aussagekraft der wichtigsten Merkmale RBD vs. PK

Eine Zuordnung zu RBD erfolgte allerdings bisher hauptsächlich anhand anderer Parameter, da diese wohl für dieses Krankheitsbild aussagekräftiger sind. Blickt man auf die wichtigsten Merkmale des Kohortenvergleichs zwischen RBD und PK, sieht man, dass erneut die beiden Merkmale der Stimmanalyse aus Abbildung 11 und 12 die beiden einflussreichsten sind. Die Abbildungen 15 und 16 zeigen die beiden Merkmale nur in umgekehrter Reihenfolge. Wie man schon in Abbildung 11 sehen kann, stellt es sich als äußerst schwierig dar, PK und Kontrollen mit diesen beiden Merkmalen zu unterscheiden. Die erneute hohe Wichtigkeit sagt uns hierbei jedoch, dass wir eine wichtige Aussage über RBD treffen können. Da aber offensichtlich mit beiden Merkmalen keine Aussage zu PK oder Kontrollen möglich ist, stellt sich die Frage, ob Merkmale der Stimmanalyse

ein geeignetes Werkzeug sind, um weitere Analysen jenseits einer bloßen Zuordnung zu einer der drei Gruppen anzustellen.

Messwerte der wertvollsten Merkmale in der Auswertung RBD vs. PK

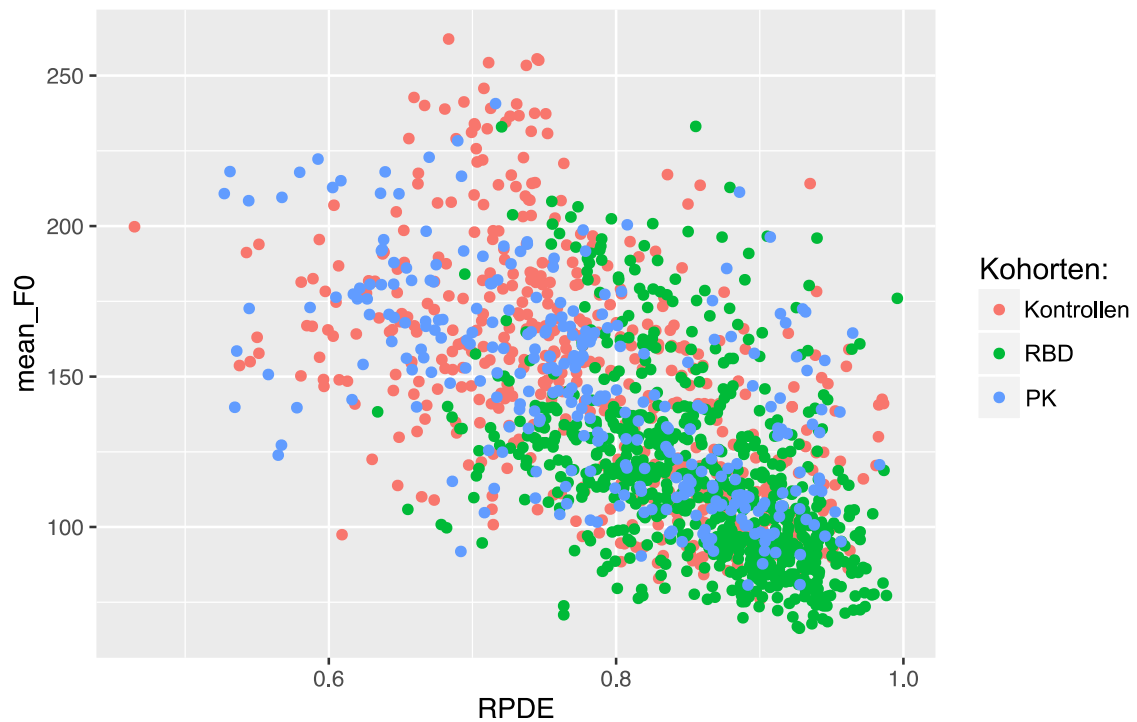


Abbildung 15: Streudiagramm der zwei wertvollsten Merkmale in der Auswertung RBD vs. PK

Abbildung 15 zeigt ein Streudiagramm, das die am stärksten gewichteten Merkmale „RPDE“ und „mean_F0“ (beides Merkmale der Stimmanalyse) für den Vergleich RBD vs. PK gegeneinander aufträgt. Jeder Punkt stellt einen Messwert dieser beiden Merkmale dar. Es werden die Messwerte aller drei Kohorten dargestellt, um mögliche Unterschiede in der Werteverteilung zu visualisieren.

Abkürzungen: PK = Parkinson-Krankheit, RBD = REM-Schlaf-Verhaltensstörung, RPDE = Recurrence Period Density Entropy.

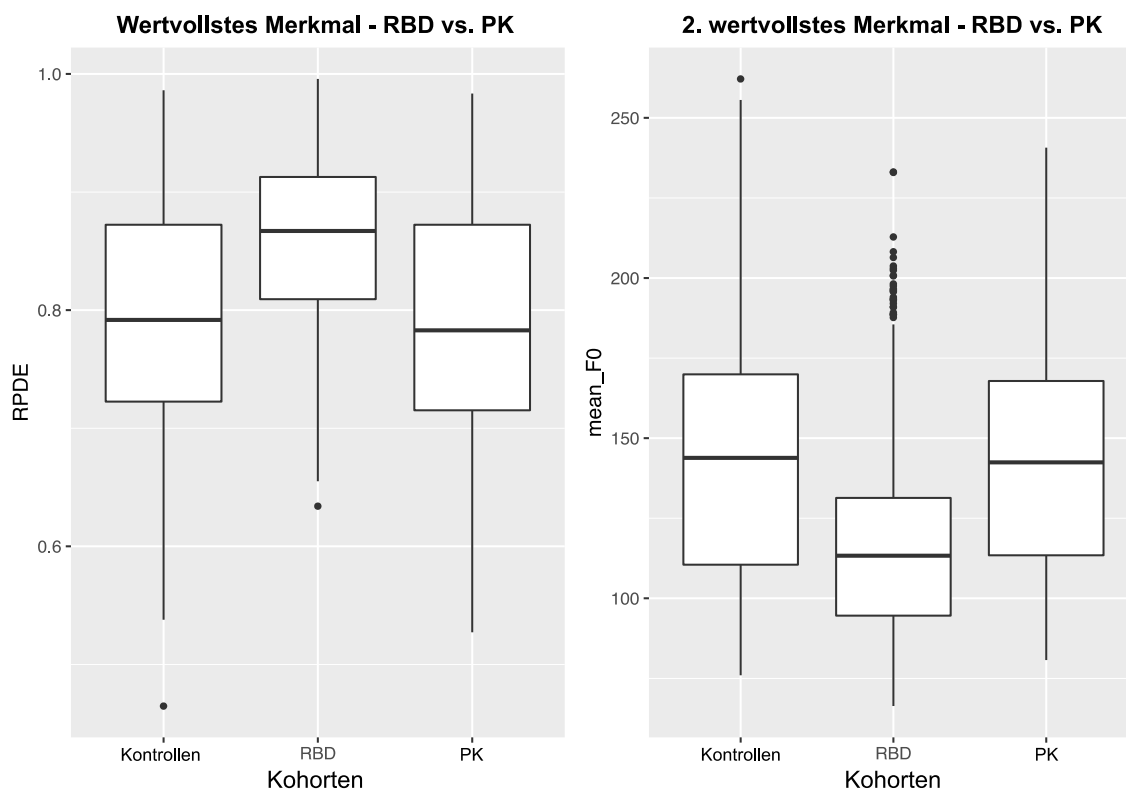


Abbildung 16: Boxplots wertvollste Merkmale RBD vs. PK

Abbildung 16 zeigt Boxplots, welche die Verteilung der Werte um die jeweiligen Mittelwerte der beiden Merkmale „RPDE“ und „mean_F0“ (beides Merkmale der Stimmanalyse) in den drei untersuchten Kohorten darstellen. Es zeigen sich deutliche Unterschiede in der Wertausprägung zwischen der RBD-Kohorte und den zwei anderen Kohorten.

Abkürzungen: RBD = REM-Schlaf-Verhaltensstörung, PK = Parkinson-Krankheit, RPDE = Recurrence Period Density Entropy.

5.4 Anteile der Tests an den wichtigsten Merkmalen mit Aussagekraft zu allen Kohorten

Vergleichen wir die Wichtigkeits--Rangfolge über alle Modelle hinweg miteinander, zeigt sich dabei in Abbildung 17, dass die aussagekräftigsten Merkmale bzw. diejenigen Merkmale, die eine Aussage zu allen drei Gruppen und nicht nur zu einer der drei Gruppen treffen, fast zu drei Vierteln aus dem Test des Haltetremors kommen. Das unterstreicht und illustriert den Informationsgehalt dieses Tests und deckt sich im Wesentlichen mit den Aussagen der Abbildungen 4, 7 und 10, in denen der Haltetremor bei der 10-fachen Kreuzvalidierung immer zu einem wesentlichen Anteil an den wichtigsten Merkmalen beteiligt ist. Auf die Stimmanalyse, den Gangtest und die Reaktionszeitmessung entfallen die übrigen aussagekräftigen Merkmale. Dies war anhand der Abbildungen 4, 7 und 10 so nicht zu erwarten, da zum Beispiel der Anteil an der Entscheidungsfindung, der auf den Gangtest entfällt, zunächst wohl eher unterschätzt worden wäre. Der Gangtest

ist sonst nur in Abbildung 7 zur Unterscheidung von PK und Kontrollen unter den Top 30 Merkmalen vertreten, scheint aber entsprechend bei den anderen Vergleichen noch wichtig genug zu sein, sodass er zu den Top 30 Merkmalen der Gesamtauswertung gehört.

Zusammensetzung der Top 30 Merkmale insgesamt

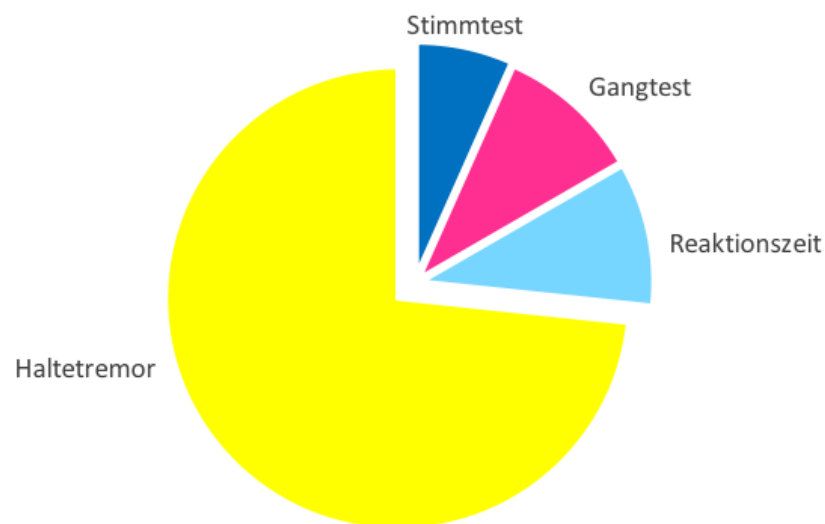


Abbildung 17: Zusammensetzung der Top 30 Merkmale insgesamt

Abbildung 17 zeigt die Herkunft der 30 wichtigsten Merkmale aus den jeweiligen Tests. Dargestellt ist somit die Herkunft der 30 Merkmale, die insgesamt den höchsten Einfluss über alle drei Gruppen PK, RBD und Kontrollen hinweg auf eine Zuordnungsentscheidung haben.

6 Diskussion

6.1 Interpretation der Ergebnisse

Wir untersuchten Kohorten von Patienten mit Parkinson-Krankheit, REM-Schlaf-Verhaltensstörung und gesunde Kontrollen, die umfangreiche Diagnostik durchlaufen haben. Bei diesen Probanden gelang es uns mittels Smartphones, die jeweiligen Patienten zu den Kohorten RBD, PK und Kontrollen zuzuordnen. Diese Zuordnung verglichen wir dann mit der wahren Gruppenzugehörigkeit. Als wahre Gruppenzugehörigkeit galt die von einem erfahrenen Kliniker vorgenommene Diagnose. Hierbei erreichten wir abhängig von der Methode der Kreuzvalidierung eine hohe statistische Sicherheit bei geeigneter Wahl der Parameter ($AUC > 0,9$). Gemessen an unserer Kohortengröße und ausgehend von der Tatsache, dass selbstlernende Algorithmen eine bessere Leistung zeigen, je größer die Stichprobe ist, konnten wir hier ein gutes Ergebnis erzielen. Weiterhin konnten wir auch die Ergebnisse der Universität Oxford bestätigen, dass die Smartphone-Sensor-Daten ebenso gut zwischen RBD-Patienten und gesunden Kontrollen wie zwischen PK-Patienten und gesunden Kontrollen unterscheiden konnten, obwohl PK-Patienten üblicherweise deutlich stärkere motorische Symptome als RBD-Patienten haben (Arora et al., 2018).

Unsere Ergebnisse zeigen eindrücklich, dass ein einmaliger Testdurchlauf dieser App wohl in näherer Zukunft nicht zu einem Ergebnis mit zufriedenstellender Zuordnungssicherheit führen wird. Die LORO-Kreuzvalidierung konnte in keinem der drei paarweisen Vergleiche überzeugende Ergebnisse liefern ($AUC < 0,76$). Das legt den Schluss nahe, dass bei einem einzelnen Testdurchlauf Störeinflüsse zu groß sind, zumindest bei einer Datenmenge wie unserer. Mittels 10-fach-KV, also der unabhängigen Betrachtung der Aufnahmen untereinander, konnte die Zuordnungssicherheit erheblich gesteigert werden ($AUC > 0,92$). Dabei ist davon auszugehen, dass mit deutlich größeren Datensätzen auch die Leistung des Modells weiter profitieren würde.

Stimmanalyse und Reaktionszeitmessung waren die wichtigsten Tests, um zwischen RBD- und Kontroll-Kohorte zu unterscheiden. Eine Unterscheidung zwischen PK und Kontrollen wurde hauptsächlich anhand von Merkmalen aus den Tests für Haltetremor und Reaktionszeitmessung sowie Gangtest und Stimmanalyse getroffen. Dass die Unterscheidung hier anhand relativ unterschiedlicher Tests getroffen wurde, überrascht nicht,

da man deutlich betontere motorische Symptome, unter anderem Tremor und Gangstörungen, in der Parkinson-Kohorte erwarten konnte. Überraschend war allerdings, dass in beiden Vergleichen die Reaktionszeitmessung ein einflussreicher Test war. Das ist ein gegensätzliches Resultat zur Vorstudie in Oxford, die die Reaktionszeitmessung in ihrer Kohorte als einen der einflussärmsten Tests beschreibt (Arora et al., 2018). Arora et al. führen das auf ein sehr durchwachsendes Engagement bei der Reaktionszeitmessung zurück. Wir konnten allerdings in den klinischen Tests beobachten, dass gerade diese Aufgabe den Probanden am meisten Spaß gemacht hat und sie besonders motiviert waren. Es bedarf weiterer Untersuchungen, ob bei entsprechender Motivation und verbesserter Anleitung die Reaktionszeitmessung in ihrer Aussagekraft verbessert werden kann.

In der Unterscheidung zwischen RBD und PK waren die wichtigsten Merkmale aus den Bereichen Stimmanalyse und Haltetremor. Ebenfalls eine Rolle spielten Merkmale aus der Analyse des Ruhetremors und der Reaktionszeitmessung.

Dass Stimmveränderungen eine Zuordnung zur RBD-Kohorte bedingen, deckt sich auch mit früheren Studien zu Stimmveränderungen bei RBD (Hlavnička et al., 2017). Insbesondere eine monotone Stimme scheint typisch für RBD Patienten, wohingegen Parkinsonpatienten zum Teil andere Stimmveränderungen aufweisen, die zum Teil auch auf dopaminerge Therapie ansprechen und daher unter Umständen nicht so konsistent über alle Messungen hinweg vorliegen (Rusz et al., 2021).

Insgesamt erhielten wir ein relativ heterogenes Feld mit wichtigen Merkmalen aus unterschiedlichen Bereichen, die bei den drei paarweisen Vergleichen offensichtlich unterschiedlich aussagekräftig waren. Der Algorithmus machte seine Zuordnung zu PK also an anderen Parametern fest als zu RBD. Das bedeutet allerdings nicht, dass es nicht möglich wäre, eine sichere Entscheidung auch anhand anderer, als der vom Algorithmus gewählten Parameter, zu treffen. Die Heterogenität der wichtigen Merkmale im Vergleich der Gruppen untereinander bedeutet, dass es nicht notwendigerweise eine kontinuierliche Veränderung der einzelnen Merkmale im Sinne einer Verschlechterung über die Zeit bei Fortschreiten der Krankheit zu geben scheint. Oder aber diese Veränderung ist sehr klein, sodass das Heranziehen anderer Parameter zuverlässiger ist, um eine Zuordnungsentscheidung zu treffen. Ein Merkmal, das eine hohe Aussagekraft zu PK hat, muss also keine hohe Aussagekraft zu RBD haben und umgekehrt. Allerdings sind Merkmale, die eine hohe Aussagekraft zu beiden Kohorten bzw. Krankheiten haben, von besonderem

Interesse, da sie geeignet sein könnten, eine Krankheitsprogression zu objektivieren. Deshalb schlossen wir eine Bewertung der Merkmale über alle Gruppen hinweg an, identifizierten so diejenigen, die eine hohe Aussagekraft zu allen drei Kohorten haben, und betrachteten diese gesondert.

Relativ gesehen hatten Merkmale aus der Haltetremor-Analyse die höchste Aussagekraft über alle drei Gruppen hinweg. Ein Ergebnis, das sich teilweise auch mit den Resultaten von Arora et al. deckt, die ebenfalls berichteten, dass Ruhe- und Haltetremor Tests eine hohe Aussagekraft bezüglich RBD und Parkinson-Krankheit haben. Diskrete Tremor-Symptome scheinen bei Patienten mit RBD häufig schon vor der Konversion zur PK und vor der Diagnose von RBD vorzuliegen. Eine retrospektive Analyse konnte zeigen, dass Menschen, die bei sich selbst oder bei denen der Hausarzt diskrete Tremor-Symptome feststellte, häufiger später eine Parkinson-Krankheit entwickelten (Schrag, Horsfall, Walters, Noyce, & Petersen, 2014). Mittels Smartphones gelang es uns nun wahrscheinlich, diskrete Tremor-Symptome, die in der klinischen Untersuchung nicht auffielen, zu objektivieren und daraus eine Aussage zum Krankheitsstatus der Patienten abzuleiten. Wir vermuten, dass repetitive Messungen mittels Smartphones der sporadischen klinischen Untersuchung in dem Erkennen subtiler Tremor-Symptome überlegen sind. Allerdings werden erst longitudinale Untersuchungen zeigen, ob sich der Tremor bei diesen Patienten später tatsächlich auch klinisch manifestiert oder eine Konversion von RBD zu PK anzeigt (Arora et al., 2018).

Weitere Hinweise darauf, dass eine Untersuchung des posturalen Tremors mittels Smartphones in der Tat ein realistisches und präzises Abbild des Krankheitsprogresses darstellt, erhielten wir aus einer weiteren Analyse der einflussreichsten Merkmale. Betrachtet man Abbildung 14, so scheinen insbesondere die beiden Merkmale „T7_PostureRootMeanSquare“ (ein Merkmal aus der Haltetremor-Analyse) und „Diff_OnOff_RT“ (ein Merkmal aus der Reaktionszeit-Analyse) eine valide Aussage bezüglich des Krankheitsstatus zuzulassen, da sich die Werte von der Kontroll-Kohorte über die RBD-Kohorte hin zur PK-Kohorte kontinuierlich verändern und damit eine mögliche kontinuierliche Verschlechterung der Symptome der Testpersonen bedeuten. Auch hier müssen longitudinale Daten zeigen, ob es sich in der Tat um valide und ausreichend sensible Parameter handelt, die auch eine Krankheitsprogression über einen längeren Zeitraum am Individuum erkennen. Reaktionszeitmessung und Haltetremor sind Tests, die in Zukunft besondere Aufmerksamkeit bei der Weiterentwicklung der App und dem Erheben von longitudinalen Daten erfahren sollten.

Dem gegenüber zeigen unsere explorativen Untersuchungen der weiteren entscheidungsrelevantesten Merkmale, dass wir den Nutzen der Stimmanalyse, zumindestens innerhalb der App, bisher unter Umständen überschätzt haben. In Abbildung 12 und 16 zeigt sich keine kontinuierliche Veränderung der beiden wichtigsten Stimmanalyse-Merkmale. Die beiden Merkmale „RPDE“ (=Recurrence period density entropy, ein Maß für Periodizität und Dynamik in der Stimme) und „meanF0“ (mittlere Grundfrequenz der Stimme) sind offensichtlich geeignet, um zwischen RBD und gesunden Kontrollen oder RBD und PK zu unterscheiden. Sie sind aber nicht oder nur sehr eingeschränkt hilfreich für eine Unterscheidung zwischen PK und Kontrollen. Genau in dieser Art scheint der Algorithmus die beiden Merkmale auch verwendet zu haben; er benutzte sie wohl als RBD-Marker. Dabei stehen diese beiden Merkmale stellvertretend für den ganzen Analyse-Cluster des Stimmtests. Eine hohe Aussagekraft besteht nur bezüglich RBD, nicht aber bezüglich der beiden anderen Kohorten. Dennoch haben diese ihre Berechtigung, um RBD-Erkrankte gegenüber Gesunden und Parkinson-Erkrankten zu identifizieren.

Worauf die paradoxe Tatsache beruht, dass Gesunde und Parkinson-Erkrankte eine sehr ähnliche Werteausprägung haben, jedoch RBD-Patienten eine deutliche Stimmveränderung zeigen, sollte in weiteren Untersuchungen geklärt werden. Auch hier lohnen sich longitudinale Datenerhebungen, um zu klären, ob die Krankheitsentwicklung hier einen parabolischen Verlauf nimmt. Denkbar ist auch ein medikamentöser Einfluss durch dopaminerge Therapien auf die Sprache von Parkinsonpatienten, insbesondere gibt es Hinweise darauf, dass die Heiserkeit unter dopaminergem Therapie bei PK abnimmt (Rusz et al., 2021). Sollte sich dies in Zukunft bestätigen, könnte hier jedoch ein Ansatzpunkt sein um einen Therapieerfolg zu bewerten und Therapien zu steuern.

In Bezug auf die Vergleichbarkeit und Gewichtung der Parameter untereinander wirft das jedoch Probleme auf, insbesondere wenn man versucht, aus den vorhandenen Daten einen Komposit-Score zu erstellen, der eine Aussage über die Symptomlast oder Fortschreiten der Erkrankung machen könnte, was wiederum erklärtes Ziel der Entwickler ist. Daher muss man sagen, dass die Analysekraft des auf den ersten Blick vielversprechenden Clusters Stimmanalyse in der App geringer ist, als das aus den vorhergehenden Analysen anzunehmen war.

Insgesamt bleibt abzuwarten, ob mit der Erhebung von longitudinalen Daten neue Erkenntnisse bezüglich der Position eines Patienten auf der Zeitachse der Erkrankung und zum Beispiel seiner Konversionswahrscheinlichkeit im nächsten Jahr gewonnen werden

können. Dabei sollten insbesondere der Haltetremor und die Reaktionszeit im Mittelpunkt der Analyse stehen.

6.2 Durchführbarkeit

Um die Validität der Ergebnisse zu gewährleisten, untersuchten wir unseren Datensatz mit unterschiedlicher Methodik. Wir verwendeten unterschiedliche Kreuzvalidierungsmethoden und untersuchten den Datensatz als Ganzes und in verschiedenen Teilmengen. Außerdem führten wir mehrere Wichtigkeits-Analysen in den einzelnen Gruppen und über alle Gruppen hinweg durch. Daher gehen wir davon aus, dass keine großen Störfaktoren unsere Ergebnisse verfälscht haben. Nichtsdestotrotz wählten wir einen Untersuchungsansatz, der weitgehend auf Kontrolle und kontrollierte Umgebungsbedingungen verzichtet, daher können Einflüsse von außen und eine Reduktion der Datenqualität nicht vermieden werden (Arora et al., 2018).

6.2.1 Studienabbrecher und Compliance

Insgesamt konnten wir bei unseren Probanden eine hohe Motivation für die Teilnahme an der Studie feststellen. Aus den Tabellen 2 und 3 ist ersichtlich, dass in der Kohorte „RBD“ 5 Probanden und in der Kohorte „gesunde Kontrollen“ 3 Probanden unzureichende Daten ablieferten und daher aus der Auswertung ausgeschlossen werden mussten. Dies geschah aufgrund unzureichender Datenmenge, mangelnder Qualität oder gänzlich fehlenden Aufzeichnungen, eventuell durch Aufzeichnungs- und Übertragungsschwierigkeiten der Applikation bedingt.

Eine hohe Compliance erreichten wir über ein integratives Studienkonzept, bei dem wir gesunde Familienmitglieder als gesunde Kontrollpersonen anwarben. Bei einem Patientenkollektiv, bei dem ein nicht unwesentlicher Teil der Probanden keine oder wenig Erfahrung in der Nutzung von Smartphones hatte, sahen wir die Notwendigkeit, die Studienadhärenz mit kreativen Mitteln zu steigern.

6.2.2 Einfluss des Alters der Probanden

Im Verlauf der Studie hatten die Probanden immer wieder mit Problemen in der Bedienung der App sowie in der allgemeinen Bedienung des Smartphones zu kämpfen. Viele der letztendlich eingeschlossenen Patienten haben schon zuvor ein Smartphone besessen,

was die Teilnahme zwar leichter machte, jedoch nicht gleichzeitig dazu führte, dass die Bedienung reibungslos funktionierte. Wir benutzten Android Smartphones, was für alle iPhone-Nutzer zwangsläufig eine Umstellung bedeutete. Insgesamt waren unsere Probanden keine „digital natives“ und insofern sehr unterschiedlich geübt im Umgang mit der Technik. Die Studienadhärenz konnte teilweise nur durch häufige Telefonate und Hilfestellungen in der Bedienung gewährleistet werden. Probanden, die vorher noch nie ein Smartphone besessen hatten, taten sich entsprechend noch schwerer. Daraus kann natürlich eine Stichprobenverzerrung (engl. „selection bias“) resultieren, sodass wir davon ausgehen müssen, dass unsere Probanden verhältnismäßig fit und kognitiv wenig eingeschränkt waren.

6.2.3 Probleme mit der Software und Hardware

Immer wieder kam es im Verlauf der Studie zu Schwierigkeiten der Datenübertragung von den Smartphones der Patienten auf die Server. Daten mussten daher zum Teil manuell verschickt werden oder gingen komplett verloren. Im Verlauf der Studie konnte durch einige Updates die Stabilität verbessert werden, allerdings war der Ablauf selten reibungslos. Verbesserungswürdig war auch die Tatsache, dass es während den Testdurchläufen zu Abstürzen der App kam, wobei der Testfortschritt bis dahin verloren ging. Testreihen mussten daraufhin wiederholt werden, da keine Daten zwischengespeichert wurden. Es ist nicht klar, ob die Abstürze auf Probleme der Hardware unserer Smartphones oder der Software zurückzuführen waren. Einige unserer Smartphones entwickelten über die Zeit, mutmaßlich durch das häufige Verschicken mit der Post, Stabilitätsprobleme sowie Probleme mit der Akku-Laufzeit. Der Stromverbrauch der App schien insgesamt sehr hoch.

In Kombination mit dem Alter der Patienten und der Tatsache, dass man diese nicht unbedingt als technikaffin beschreiben kann, potenzierten sich solche Probleme natürlich und führten dazu, dass einige Patienten die Studie mehrfach abbrechen wollten. Nur durch viel Überzeugungsarbeit konnte die weitere Teilnahme gesichert werden. Allerdings entschieden aufgrund der Erfahrungen nicht wenige, dass sie für eine ursprünglich geplante Follow-Up Untersuchung nicht zur Verfügung stehen.

Im Gegensatz zu Vorstudien durften wir, aufgrund von Datenschutzbedenken der hiesigen Ethikkommission, nicht auf die persönlichen Smartphones der Probanden zurückgreifen, um die App dort zu installieren. In Bezug auf die Compliance der Patienten hatte

das einige Nachteile. Probanden mussten so, wenn sie unterwegs waren, immer zwei Smartphones mit sich führen. Dieser Mehraufwand führte dazu, dass weniger Tests durchgeführt wurden und der Alltag nicht in Gänze abgebildet wurde, weil einige Probanden dann einfach keine Tests durchführten.

Weiterhin waren die Probanden im Umgang mit unseren Smartphones nicht geübt, sodass sie sich erst im neuen Betriebssystem und Layout zurechtfinden mussten und die Bedienung nicht so selbstverständlich war wie mit den eigenen Smartphones.

6.3 Eigene Limitationen und Optimierungsmethoden

6.3.1 Einfluss des Geschlechts der Probanden

Bei Betrachten der demographischen Daten in Tabelle 1 fällt auf, dass auch in unserer Stichprobe Frauen in der Kohorte der RBD bzw. Parkinson Erkrankten deutlich unterrepräsentiert sind. Andererseits sind 21 von 34 Kontrollpersonen weiblich. Dies bildet die tatsächlichen Unterschiede der Prävalenz nur teilweise ab, insbesondere in der Kohorte der Patienten mit idiopathischem Parkinson-Syndrom, denn hier konnte ein signifikanter Prävalenzunterschied bisher nur in der Kohorte der 50- bis 59-jährigen nachgewiesen werden (Pringsheim et al., 2014). Die REM-Schlaf-Verhaltensstörung betrifft jedoch tatsächlich hauptsächlich Männer, beziehungsweise wird bei Männern häufiger diagnostiziert (Olson et al., 2000). Insofern war in dieser Kohorte eine hauptsächlich männliche Population erwartbar.

Wir schätzen allerdings den Einfluss dieser Diskrepanz auf die Ergebnisse der Studie als verhältnismäßig gering ein, da eine große Zuordnungssicherheit über alle drei Vergleiche erreicht werden konnte und zumindest die Gruppen RBD und PK in ihrem Geschlechterverhältnis sehr ähnlich waren. Ein Einfluss ist aber nicht auszuschließen.

Die vermehrte Rekrutierung der Partnerinnen für die Gruppe der gesunden Kontrollpersonen war mit ursächlich für das ungleiche Geschlechterverhältnis. Dadurch konnten wir aber Studienabbrüchen vorbeugen, da die Tests mit einem Partner zur gleichen Zeit in der Familie durchgeführt werden konnten.

6.3.2 Limitation der Applikation und des Studiendesigns

In Studien konnten mit einer Genauigkeit von ca. 99% Patienten mit idiopathischem Parkinson-Syndrom von gesunden Kontrollpersonen unterschieden werden, wobei nur Sprachaufnahmen der Probanden verwendet wurden (Tsanas, Little, McSharry, Spielman, & Ramig, 2012). Studien wie diese setzten bisher auf ein hohes Maß an Umgebungskontrolle, um Störfaktoren und Einflüsse von außen auszuschalten. In einem Studiendesign wie dem unseren wird bewusst auf diese Umgebungskontrolle verzichtet, um im Gegenzug dafür die Möglichkeit einer kontinuierlicheren und breiteren Datenerhebung zu schaffen, ohne die Probanden in ihrem Alltag einzuschränken. Dabei ist eine verminderte Datenqualität unausweichlich (Arora et al., 2018). Die Probanden wurden angehalten, Störeinflüsse von außen so weit wie möglich gering zu halten, dennoch kann eine Vielzahl von Einflüssen nicht ausgeschlossen werden.

6.3.2.1 Äußere Einflüsse

Äußere Einflüsse stellen bei einem Studiendesign wie dem unseren immer eine mögliche Fehlerquelle bzw. einen Störfaktor dar. Die Spanne der denkbaren Einflüsse ist breit und ist maßgeblich abhängig davon, welchen Test wir betrachten.

Bei der Stimmanalyse müssen Umgebungsgeräusche und Mikrofonpositionierung beachtet werden. Wir hielten die Patienten an, sich nach Möglichkeit einen ruhigen und an Umgebungsgeräuschen armen Ort zu suchen, an dem sie ihren Test durchführen sollten. Dennoch ist vorstellbar, dass einige Patienten ihre Aufzeichnungen an wechselnden Orten oder Orten mit Hintergrundgeräuschen durchführten.

Bei der Gleichgewichtstestung und der Gangtestung könnten Smartphone-Positionierung, Taschengröße und Ausrichtung des Smartphones eine mögliche Störgröße darstellen und die Daten des Akzelerometers beeinflussen (Arora et al., 2018). Außerdem ist nicht immer gewährleistet, dass die Gehstrecke von 20 Metern am Stück gegangen werden kann, sodass manche Patienten mehr Wendungen absolvierten als andere.

Insgesamt war auch die Auswahl an Tests und Merkmalen dadurch limitiert, dass ein einzelnes Gerät zur Datenerhebung genutzt wurde (Arora et al., 2018). Zudem konnten wir auf Grund fehlender Kontrollmechanismen nicht genau überprüfen, wie genau sich die Probanden an das Testprotokoll hielten. Auf Grund der Effektgröße und Genauigkeit in der Zuordnung gehen wir aber davon aus, dass wir verlässliche Ergebnisse produzieren konnten.

6.3.2.2 Nebendiagnosen

Es ist denkbar, dass weitere Erkrankungen der Patienten die Messdaten beeinflusst haben. Insbesondere orthopädische Leiden könnten einen Einfluss auf die motorischen Daten gehabt haben. Es berichten in einer Metanalyse beispielsweise rund 25% der Befragten über 60 Jahre über chronische untere Rückenschmerzen. Daher ist von einer erheblichen Krankheitslast und damit einhergehenden motorischen Einschränkungen auch in unserer Studienpopulation auszugehen. Es zeigte sich auch, dass die Krankheitslast hier bei Frauen ca. 50% höher ist als bei Männern (Meucci, Fassa, & Faria, 2015). Auch ein Einfluss anderer internistischer oder neurologischer Grunderkrankungen ist nicht auszuschließen. Da sich die gesunden Kontrollpersonen hinsichtlich des Alters kaum von unseren beiden anderen Testkohorten unterscheiden, sollten sich diese Einflüsse in den verschiedenen Kohorten ausbalancieren, jedoch ist bei einer Stichprobengröße wie unserer nicht mit letzter Sicherheit zu sagen, ob die Einflüsse insbesondere des Geschlechts zu vernachlässigen sind.

6.3.2.3 Medikamente und Wirkungsfluktuationen

Eine weitere mögliche Fehlerquelle liegt in dem Einfluss der Medikamente auf die Testperformance insbesondere der PK-Kohorte. Möglicherweise fühlten sich medikamentös schlechter eingestellte Patienten überhaupt nicht in der Lage an unserer Studie teilzunehmen, sodass hier eine Stichprobenverzerrung (engl. „selection bias“) zugrunde liegen könnte.

Außerdem könnten die Testzeitpunkte von den Probanden so gewählt worden sein, dass sie eher in den On-Perioden des Tages lagen, da ihnen die Testdurchführung dann leichter gefallen ist. In dieser Hinsicht könnten wir die tatsächliche Effektgröße unterschätzt haben (Arora et al., 2018). In Ermangelung von Kontrollbeobachtungen, wie Videoaufzeichnungen oder einem Symptomtagebuch, wiesen wir die Patienten an, die Tests möglichst immer zur selben Tageszeit und in homogenen Intervallen durchzuführen, um so ein der tatsächlichen Konstitution der Probanden entsprechendes Abbild erheben zu können.

Insgesamt könnte hier die Erfassung von Nebendiagnosen und Zeitpunkts- und Substanzerfassung der Medikamenteneinnahme die Datenqualität verbessern, jedoch würde das die Datenauswertung auch weiter verkomplizieren und die Handhabung der App für die Studienteilnehmer erschweren.

6.3.3 Statistische und methodische Limitationen

Generell profitieren statistische „Machine-Learning“-Algorithmen von einer großen Datenmenge. Es ist davon auszugehen, dass sich sowohl die Unterscheidungsgenauigkeit der 10-fach-KV als auch der LORO-KV signifikant verbessern wird, je mehr Probanden und Daten eingeschlossen werden können. Die LORO-KV schnitt in unserer Analyse deutlich schlechter ab als die 10-fach-KV. Dies ist insofern bedauerlich, da eine LORO-KV-Analyse dem klinisch bedeutsamen Szenario der Untersuchung einer bis dato unbekannt Person mit einer vortrainierten Applikation am nächsten kommt.

Es sind eine Menge Einflüsse denkbar, die auf die Datenqualität eingewirkt haben könnten. Wir testeten Probanden über eine Woche, jedoch könnte ein noch längerer Testzeitraum den Einfluss von Störfaktoren, wie zum Beispiel Umgebungsgeräuschen, begrenzen und die Wahrscheinlichkeit erhöhen, qualitativ hochwertige Aufzeichnungen zu erfassen. Patienten steuerten zwar im Mittel 16,3-19,9 Aufnahmen bei, damit lagen sie aber deutlich unter den angestrebten 28 Aufnahmen pro Patienten pro Woche. Auch hier könnte ein längerer Testzeitraum dazu führen, dass die Zielvorgabe erreicht wird.

Insgesamt ist es nicht möglich, bei komplexen „Machine-Learning“-Algorithmen auszuschließen, dass diese Vorhersagen aufgrund von Merkmalen machen, die in keiner Relation zum Krankheitsstatus der Patienten stehen. Den genauen Einfluss dieser Störfaktoren zu quantifizieren, ist aufgrund der hohen Komplexität der Algorithmen nicht möglich (Arora et al., 2018; Little et al., 2017). Wir versuchten, durch das Heranziehen unterschiedlicher Kreuzvalidierungsverfahren dieses Problem zu minimieren, mussten jedoch feststellen, dass die Unterscheidungsgenauigkeit der LORO-KV deutlich niedriger war als die der 10-fach-KV.

Des Weiteren spielt die Auswahl des Algorithmus für die Wichtigkeits-Bewertung eine wichtige Rolle, da verschiedene Algorithmen auch verschiedene Kriterien anlegen, um die Wichtigkeit der verschiedenen Merkmale zu bewerten. Je nachdem welcher Algorithmus also gewählt wurde, kann sich das Ergebnis unterscheiden. Die in Abbildung 4,7 und 10 gezeigten Kreisdiagramme erlauben zwar eine Einschätzung, wie stark der Einfluss der einzelnen Tests auf die Zuordnungsentscheidung war, eine präzisere Angabe, zu wieviel Prozent ein einzelnes Merkmal die Entscheidung im Mittel beeinflusst, bleibt diese Analyse jedoch schuldig. Um dennoch genauere Aussagen über den Einfluss der einzelnen Tests und Merkmale machen zu können, bewerteten wir ihren Einfluss über alle drei paarweisen Vergleiche hinweg, wobei wir eine simple Bewertung verwendeten, die sich

auf die Summe der einzelnen Wichtigkeits-Bewertungen stützt. Dadurch konnten wir unsere Einschätzung zur Wichtigkeit der einzelnen Merkmale zwar verbessern, eine genaue Quantifizierung blieb aber unmöglich. Insofern besteht das Problem, dass die Datenanalyse zu komplex war, um sie eins zu eins auf klinische Befunde zu übertragen. Dennoch konnten wir hier die Einschätzung, aufgrund welcher Tests bzw. Merkmale eine Entscheidung getroffen wurde, deutlich präzisieren.

Wir waren in unserer Analyse auf klinische Beobachtungen angewiesen, um die zugrundeliegende Wirklichkeit der Gruppenzugehörigkeit festzulegen. Unsere Patienten wurden ausschließlich im Universitätsklinikum Marburg rekrutiert, insofern konnten hohe diagnostische Ansprüche gewährleistet werden und wir sind zuversichtlich, dass unsere Diagnosen auch nach neuropathologischer Überprüfung Bestand haben. Dennoch muss beachtet werden, dass bis zu ca. 25% der initial mit einer Parkinson-Krankheit diagnostizierten Patienten, wenn diese Diagnose nicht von einem auf Bewegungsstörungen spezialisierten Neurologen gestellt wurde, nach neuropathologischer Überprüfung eine alternative Diagnose erhalten (Joutsa, Gardberg, Røyttä, & Kaasinen, 2014). Fehldiagnosen sind daher nicht auszuschließen und, um die Präzision weiter zu erhöhen, sind Langzeitstudien mit anschließender neuropathologischer Untersuchung der Patienten post mortem nötig.

6.4 Bedeutung der Ergebnisse und Fazit

Diese Studie konnte zeigen, dass die Applikation (kurz: App) „HopkinsPD“ unter Nutzung von gängigen Smartphones und den dort integrierten IMUs eine Unterscheidung zwischen gesunden Kontrollpersonen, PK-Patienten und Patienten mit RBD mit hoher Sicherheit ermöglicht. Es konnten die einflussreichsten und wichtigsten Merkmale identifiziert und ihr spezifischer Beitrag zur Entscheidungsfindung genauer charakterisiert werden. Dabei zeigte sich, dass in früheren Studien die tatsächliche Aussagekraft einiger Tests beziehungsweise Merkmale teilweise überschätzt wurde, da eine direkte Übertragung auf die Klinik und den tatsächlichen Zustand des Patienten nur eingeschränkt möglich ist. Wir konnten hingegen andere Merkmale und Tests identifizieren, die sich als vielversprechende Untersuchungsansätze herausstellten, um Krankheitsprogression oder Symptomverschlechterung kontinuierlich zu beobachten. Diese sollten in Zukunft bei der Verbesserung und der Feineinstellung der Algorithmen in den Fokus gerückt werden.

Auf dem Weg hin zu einem verlässlichen und objektiven Messinstrument, dass die individuelle Symptomlast und das Konversionsrisiko in ein späteres Krankheitsstadium bestimmt, gibt es noch einige Hürden zu überwinden. Diese Studie konnte einerseits unterstreichen, dass der verfolgte Ansatz sehr vielversprechend ist, und andererseits weitere Entwicklungsansätze zur Verfeinerung der Analysemethoden darlegen.

Letztendlich wird es noch weitere Jahre und die Erhebung großer Datenmengen erfordern, um eine automatisierte und individuell nutzbare Applikation zu entwickeln, die auch robust in der Abgrenzung gegenüber Erkrankungen ist, die ein ähnliches klinisches Bild produzieren wie die PK und RBD.

6.5 Ausblick

In zukünftigen Generationen wird die Nutzung und das Erfassen von Gesundheitsdaten mittels Smartphones mutmaßlich eine immer größere Rolle spielen. Auch der Umgang mit digitalisierter Infrastruktur wird immer natürlicher und selbstverständlicher werden. Insofern werden einige der Probleme, die wir bei der Durchführung dieser Studie hatten, von selbst verschwinden. Dennoch wird es eine große Herausforderung werden, eine Applikation wie diese zur Marktreife zu bringen. Es sind Optimierungen am Algorithmus nötig, um die Zuordnungssicherheit zu verbessern und die direkte Übertragbarkeit der Ergebnisse von Smartphone-Messungen auf die Klinik der Patienten zu gewährleisten. Die Datenbasis unserer Untersuchungen muss, um klinischen Ansprüchen zu genügen, noch deutlich breiter werden. Insbesondere sollten andere Bewegungsstörungen als Vergleichsgruppen eingeschlossen werden, da diese unter Umständen schwieriger von einer Parkinson-Krankheit zu differenzieren sind (Arora et al., 2018).

Von großem klinischem Nutzen wäre auch ein mit aktuell verwendeten Messinstrumenten der Krankheitsausprägung und Symptomprogression vergleichbarer Score. Gelänge es, die Symptomschwere mit Smartphones genau zu quantifizieren, könnte dies beispielsweise mit dem MDS-UPDRS verglichen werden. Eine quantitative Erfassung der Symptomschwere - zu jeder Tageszeit - mittels Bewegungssensoren würde ein neues Niveau an Genauigkeit und Objektivität in der Untersuchung von PK-Symptomen bedeuten. Dies würde ein kontinuierliches und patientenkontrolliertes Therapiemonitoring und dadurch verbesserte und genauere Medikamenten- und Dosisanpassungen ermöglichen. Gegen-

über der sporadischen Untersuchung durch niedergelassene Ärzte oder Spezialambulanzen wäre dies ein immenser Vorteil, der die betreuenden Ärzte in der Beurteilung der Patienten und bei Therapieentscheidungen unterstützen könnte.

Eine alternde Bevölkerung wird langfristig eine Nachfrage nach mehr Ärzten und intensiverer Betreuung schaffen. Die häusliche Beurteilung mittels digitaler Hilfsmittel, das sog. „Home-Assessment“, kann einen Beitrag leisten, diesen Druck auf das Gesundheitssystem etwas zu lindern. In Zukunft werden Applikationen eine bessere und auch schnellere Einschätzung der Symptomlast von Patienten ermöglichen. Diese Studie ist ein weiterer Schritt hin zu einer automatisierten objektiven Einschätzung von PK- und RBD-Patienten. Damit verbindet sich die Hoffnung, zukünftig eine bessere Beurteilung vornehmen und aufgrund der ermittelten Daten verbesserte Therapieentscheidungen treffen und optimieren zu können. Ebenfalls würde so der Erforschung und Evaluation von neuroprotektiven Therapien und deren Wirksamkeit Vorschub geleistet.

7 Zusammenfassung

Zielsetzung

Wir versuchten mittels gängigen Smartphones motorische Symptome von Parkinson (PK)-Patienten, Patienten mit REM-Schlaf-Verhaltensstörung (RBD) und gesunden Kontrollen zu erfassen und daraus abzuleiten, welcher der drei Gruppen diese angehörten und wie stark die Ausprägung der Symptome war.

Methodik

Insgesamt 43 RBD-, 19 PK- und 34 Kontroll-Probanden führten über eine Woche viermal täglich sieben Tests durch, die Stimme, Gleichgewicht, Gang, Geschicklichkeit, Reaktionszeit, Ruhetremor und Haltetremor aufzeichneten. Die Probanden führten die Tests einmal unter Aufsicht in der Klinik, dann sieben Tage lang mehrmals täglich unter unkontrollierten Bedingungen zu Hause durch. Die PK- und RBD-Patienten wurden in unseren Spezialambulanzen rekrutiert. Wir versuchten nur mittels der Smartphone Aufzeichnungen zuzuordnen, welcher der drei Gruppen die Patienten angehörten. Weiterhin versuchten wir, die für die Unterscheidung geeignetsten Tests und Merkmale zu identifizieren und genauer zu charakterisieren.

Ergebnisse

Es konnte mit guter statistischer Sicherheit zwischen den verschiedenen Gruppen unterschieden werden. Für die drei paarweisen Vergleiche PK versus Kontrollen, RBD versus Kontrollen und RBD versus PK lagen die AUCs zwischen 0,82 - 0,95 bei 10 - 50 inkludierten Merkmalen. Über alle Vergleiche hinweg waren Haltetremor, Ruhetremor, Stimmanalyse und Reaktionszeit die wichtigsten Tests. Dabei konnten wir zeigen, dass sich insbesondere Haltetremor und Reaktionszeit mit Fortschreiten der Krankheit kontinuierlich verändern.

Fazit

Smartphones sind in der Lage, subtile Veränderungen in der Motorik wahrzunehmen und diese objektiv zu quantifizieren. Sie können mit Hilfe von Apps und statistischen maschinellen Lernens (engl.: „Machine-Learning“) zwischen RBD, PK und gesunden Kontrollen unterscheiden. Um nur mit Smartphones ein präzises Abbild des Krankheitsstatus des

Patienten zu erstellen, sind weitere Forschung und die Verfeinerung der Analysealgorithmen nötig. Zukünftig wird die Entwicklung eines Smartphone-basierten Hilfsmittels zur Messung der Symptomschwere von Patienten sehr wahrscheinlich erfolgreich sein und ein besseres Therapiemonitoring der Patienten ermöglichen.

8 Summary

Objective

Our goal was to monitor motor symptoms of patients with Parkinson's-Disease (PK), patients with REM-sleep-behavior-disorder (RBD) and healthy controls with conventional Smartphones and derive hereby to which of the three groups the monitored patient belongs and characterize the specific manifestation of the symptoms.

Methods

Overall 43 RBD-, 19 PK- and 34 control-patients were conducting seven tests four times a day over a duration of one week to measure voice, balance, gait, dexterity (finger tapping), reaction time, rest tremor and postural tremor. The patients were conducting the seven tests once under supervision in the clinic and afterwards at home under uncontrolled conditions. The Parkinson- and RBD-patients were recruited in our specialized ambulances. We now tried to assign the patients to their cohort, using only the smartphone recordings. Furthermore, we tried to identify and characterize the most discriminative and salient features and tests for the assignment to the groups.

Results

We were able to discriminate the three groups with good statistical certainty. For the three pairwise comparisons PK versus controls, RBD versus controls and PK versus RBD the AUCs were between 0,82 - 0,95 for 10 - 50 incorporated features. In an overall comparison postural tremor, rest tremor, voice and reaction time were the most important tests. We were able to show that especially postural tremor and reaction time are continuously changing with disease progression.

Conclusion

Smartphones are able to detect subtle changes in motor function and quantify them objectively. They are able to distinguish between PK, RBD and healthy controls using Apps and statistical machine learning. To get a precise picture of the patient and his disease-status using only smartphones, further research and development of algorithms is needed. In the future the development of smartphone-based tools to measure the patient's symptoms will most likely be successful and improve the management of therapies.

9 Literaturverzeichnis

- American Academy of Sleep Medicine. (2005). *International Classification of Sleep Disorders: Diagnostic and Coding Manual. (ICSD-2). International Classification of Sleep Disorders.*
- Anaclet, C., Lin, J.-S., Vetrivelan, R., Krenzer, M., Vong, L., Fuller, P. M., & Lu, J. (2012). Identification and Characterization of a Sleep-Active Cell Group in the Rostral Medullary Brainstem. *Journal of Neuroscience*, *32*(50), 17970–17976. <https://doi.org/10.1523/JNEUROSCI.0620-12.2012>
- Arora, S., Venkataraman, V., Zhan, A., Donohue, S., Biglan, K. M., Dorsey, E. R., & Little, M. A. (2015). Detecting and monitoring the symptoms of Parkinson's disease using smartphones: A pilot study. *Parkinsonism and Related Disorders*, *21*(6), 650–653. <https://doi.org/10.1016/j.parkreldis.2015.02.026>
- Arora, Siddharth, Baig, F., Lo, C., Barber, T. R., Lawton, M. A., Zhan, A., ... Hu, M. T. (2018). Smartphone motor testing to distinguish idiopathic REM sleep behavior disorder, controls, and PD. *Neurology*, *91*(16), e1528–e1538. <https://doi.org/10.1212/WNL.0000000000006366>
- Barrantes, S., Sánchez Egea, A. J., González Rojas, H. A., Martí, M. J., Compta, Y., Valldeoriola, F., ... Valls-Solè, J. (2017). Differential diagnosis between Parkinson's disease and essential tremor using the smartphone's accelerometer. *PLOS ONE*, *12*(8), e0183843. <https://doi.org/10.1371/journal.pone.0183843>
- Bauso, D. J., Tartari, J. P., Stefani, C. V., Rojas, J. I., Giunta, D. H., & Cristiano, E. (2012). Incidence and prevalence of Parkinson's disease in Buenos Aires City, Argentina. *European Journal of Neurology*, *19*(8), 1108–1113. <https://doi.org/10.1111/j.1468-1331.2012.03683.x>
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning.* (M. Jordan, J. Kleinberg, & B. Schölkopf, Eds.). New York: Springer.
- Boeve, B. F. (2013). Idiopathic REM sleep behaviour disorder in the development of Parkinson's disease. *The Lancet Neurology*, *12*(5), 469–482. [https://doi.org/10.1016/S1474-4422\(13\)70054-1](https://doi.org/10.1016/S1474-4422(13)70054-1)
- Braak, H., Tredici, K. Del, Rüb, U., de Vos, R. A. ., Jansen Steur, E. N. ., & Braak, E. (2003). Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiology of Aging*, *24*(2), 197–211. [https://doi.org/10.1016/S0197-4580\(02\)00065-9](https://doi.org/10.1016/S0197-4580(02)00065-9)
- Breiman, L. (2001). Random Forests. *Machine Learning*, *45*, 5–32. <https://doi.org/https://doi.org/10.1023/A:1010933404324>
- Dauvilliers, Y., Schenck, C. H., Postuma, R. B., Iranzo, A., Luppi, P.-H., Plazzi, G., ... Boeve, B. (2018). REM sleep behaviour disorder. *Nature Reviews Disease Primers*, *4*(1), 19. <https://doi.org/10.1038/s41572-018-0016-5>
- Deeb, J., Shah, M., Muhammed, N., Gunasekera, R., Gannon, K., Findley, L. J., & Hawkes, C. H. (2010). A basic smell test is as sensitive as a dopamine transporter scan: comparison of olfaction, taste and DaTSCAN in the diagnosis of Parkinson's disease. *QJM*, *103*(12), 941–952. <https://doi.org/10.1093/qjmed/hcq142>
- DGN. (2016). S3-Leitlinie Idiopathisches Parkinson-Syndrom. *Leitlinien Für Diagnostik Und Therapie in Der Neurologie*, 312. Retrieved from www.dgn.org
- Dorsey, E., Constantinescu, R., Thompson, J., Biglan, K., Holloway, R., & Kiebertz, K. (2007). Projected number of people with Parkinson disease in the most populous nations,

- 2005 through 2030. *Neurology*, 68, 384–386.
<https://doi.org/10.1212/01.wnl.0000271777.50910.73>
- Doty, R. L., Shaman, P., Applebaum, S. L., Giberson, R., Siksorski, L., & Rosenberg, L. (1984). Smell Identification Ability: Changes with Age. *Science*, 226(4681), 1441–1443.
<https://doi.org/10.1126/science.6505700>
- Driver, J. A., Logroscino, G., Gaziano, J. M., & Kurth, T. (2009). Incidence and remaining lifetime risk of Parkinson disease in advanced age. *Neurology*, 72(5), 432–438.
<https://doi.org/10.1212/01.wnl.0000341769.50075.bb>
- Dugger, B. N., Boeve, B. F., Murray, M. E., Parisi, J. E., Fujishiro, H., Dickson, D. W., & Ferman, T. J. (2012). Rapid eye movement sleep behavior disorder and subtypes in autopsy-confirmed dementia with Lewy bodies. *Movement Disorders: Official Journal of the Movement Disorder Society*, 27(1), 72–78. <https://doi.org/10.1002/mds.24003>
- Dugger, B. N., Murray, M. E., Boeve, B. F., Parisi, J. E., Benarroch, E. E., Ferman, T. J., & Dickson, D. W. (2012). Neuropathological analysis of brainstem cholinergic and catecholaminergic nuclei in relation to rapid eye movement (REM) sleep behaviour disorder. *Neuropathology and Applied Neurobiology*, 38(2), 142–152.
<https://doi.org/10.1111/j.1365-2990.2011.01203.x>
- Ehringer, H., & Hornykiewicz, O. (1960). Verteilung Von Noradrenalin Und Dopamin (3-Hydroxytyramin) Im Gehirn Des Menschen Und Ihr Verhalten Bei Erkrankungen Des Extrapyramidalen Systems. *Klinische Wochenschrift*, 38(24), 1236–1239.
<https://doi.org/10.1007/BF01485901>
- Fargo, K., & Bleiler, L. (2014). 2014 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, 10(2), e47–e92. <https://doi.org/10.1016/j.jalz.2014.02.001>
- Fraigne, J. J., Torontali, Z. A., Snow, M. B., & Peever, J. H. (2015). REM Sleep at its Core - Circuits, Neurotransmitters, and Pathophysiology. *Frontiers in Neurology*, 6(MAY), 1–9.
<https://doi.org/10.3389/fneur.2015.00123>
- Galbiati, A., Verga, L., Giora, E., Zucconi, M., & Ferini-Strambi, L. (2019). The risk of neurodegeneration in REM sleep behavior disorder: A systematic review and meta-analysis of longitudinal studies. *Sleep Medicine Reviews*, 43, 37–46.
<https://doi.org/10.1016/j.smrv.2018.09.008>
- Gibb, W. R. (1988). Accuracy in the clinical diagnosis of parkinsonian syndromes. *Postgraduate Medical Journal*, 64(751), 345–351. <https://doi.org/10.1136/pgmj.64.751.345>
- Gustavsson, A., Svensson, M., Jacobi, F., Allgulander, C., Alonso, J., Beghi, E., ... Olesen, J. (2011). Cost of disorders of the brain in Europe 2010. *European Neuropsychopharmacology*, 21(10), 718–779. <https://doi.org/10.1016/j.euroneuro.2011.08.008>
- Hawkes, C. (2003). Olfaction in neurodegenerative disorder. *Movement Disorders*, 18(4), 364–372. <https://doi.org/https://doi.org/10.1002/mds.10379>
- Hawkes, C H, Shephard, B. C., & Daniel, S. E. (1997). Olfactory dysfunction in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 62(5), 436–446.
<https://doi.org/10.1136/jnnp.62.5.436>
- Hawkes, Christopher H, Tredici, K. Del, & Braak, H. (2010). A timeline for Parkinson's disease. *Parkinsonism and Related Disorders*, 16, 79–84.
<https://doi.org/10.1016/j.parkreldis.2009.08.007>
- Hlavnička, J., Čmejla, R., Tykalová, T., Šonka, K., Růžička, E., & Ruzs, J. (2017). Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder. *Scientific Reports*, 7(1), 12.

- <https://doi.org/10.1038/s41598-017-00047-5>
- Holmes, R. J., Oates, J. M., Phyland, D. J., & Hughes, A. J. (2000). Voice characteristics in the progression of Parkinson's disease. *International Journal of Language and Communication Disorders*, *35*(3), 407–418. <https://doi.org/10.1080/136828200410654>
- Hummel, T., Sekinger, B., Wolf, S. R., Pauli, E., & Kobal, G. (1997). 'Sniffin' Sticks': Olfactory Performance Assessed by the Combined Testing of Odour Identification, Odor Discrimination and Olfactory Threshold. *Chemical Senses*, *22*(1), 39–52. <https://doi.org/10.1093/chemse/22.1.39>
- Jennings, D., Siderowf, A., Stern, M., Seibyl, J., Eberly, S., Oakes, D., & Marek, K. (2017). Conversion to Parkinson Disease in the PARS Hyposmic and Dopamine Transporter–Deficit Prodromal Cohort. *JAMA Neurology*, *74*(8), 933. <https://doi.org/10.1001/jamaneurol.2017.0985>
- Joutsa, J., Gardberg, M., Røyttä, M., & Kaasinen, V. (2014). Diagnostic accuracy of parkinsonism syndromes by general neurologists. *Parkinsonism & Related Disorders*, *20*(8), 840–844. <https://doi.org/10.1016/j.parkreldis.2014.04.019>
- Kalia, L. V., & Lang, A. E. (2015). Parkinson's disease. *The Lancet*, *386*(9996), 896–912. [https://doi.org/10.1016/S0140-6736\(14\)61393-3](https://doi.org/10.1016/S0140-6736(14)61393-3)
- Krenzer, M., Anacleit, C., Vetrivelan, R., Wang, N., Vong, L., Lowell, B. B., ... Lu, J. (2011). Brainstem and Spinal Cord Circuitry Regulating REM Sleep and Muscle Atonia. *PLoS ONE*, *6*(10), e24998. <https://doi.org/10.1371/journal.pone.0024998>
- 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. <https://doi.org/10.1002/mds.26693>
- Levin, J., & Oertel, W. (2021). Proteinbezogene Therapien neurodegenerativer Erkrankungen. *Der Nervenarzt*, *92*(12), 1217–1218. <https://doi.org/10.1007/s00115-021-01218-6>
- Lindstrom, F., Ohlsson, A.-C., Sjöholm, J., & Wayne, K. P. (2010). Mean F0 Values Obtained Through Standard Phrase Pronunciation Compared With Values Obtained From the Normal Work Environment: A Study on Teacher and Child Voices Performed in a Preschool Environment. *Journal of Voice*, *24*(3), 319–323. <https://doi.org/10.1016/j.jvoice.2008.10.006>
- Little, M. A., McSharry, P. E., Roberts, S. J., Costello, D. A. E., & Moroz, I. M. (2007). Exploiting Nonlinear Recurrence and Fractal Scaling Properties for Voice Disorder Detection. *BioMedical Engineering OnLine*, *6*(1), 23. <https://doi.org/10.1186/1475-925X-6-23>
- Little, M. A., Varoquaux, G., Saeb, S., Lonini, L., Jayaraman, A., Mohr, D. C., & Kording, K. P. (2017). Using and understanding cross-validation strategies. Perspectives on Saeb et al. *GigaScience*, *6*, 1–6. <https://doi.org/10.1093/gigascience/gix020>
- Meucci, R. D., Fassa, A. G., & Faria, N. M. X. (2015). Prevalence of chronic low back pain: systematic review. *Revista de Saúde Pública*, *49*, 1. <https://doi.org/10.1590/S0034-8910.2015049005874>
- Miglis, M. G., Adler, C. H., Antelmi, E., Arnaldi, D., Baldelli, L., Boeve, B. F., ... Oertel, W. H. (2021). Biomarkers of conversion to α -synucleinopathy in isolated rapid-eye-movement sleep behaviour disorder. *The Lancet Neurology*, *20*(8), 671–684. [https://doi.org/10.1016/S1474-4422\(21\)00176-9](https://doi.org/10.1016/S1474-4422(21)00176-9)
- Nalls, M. A., Pankratz, N., Lill, C. M., Do, C. B., Hernandez, D. G., Saad, M., ... Singleton,

- A. B. (2014). Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nature Genetics*, 46(9), 989–993. <https://doi.org/10.1038/ng.3043>
- Noyce, A. J., Bestwick, J. P., Silveira-Moriyama, L., Hawkes, C. H., Giovannoni, G., Lees, A. J., & Schrag, A. (2012). Meta-analysis of early nonmotor features and risk factors for Parkinson disease. *Annals of Neurology*, 72(6), 893–901. <https://doi.org/10.1002/ana.23687>
- Oertel, W., Deuschl, G., & Poewe, W. (2012). *Parkinson-Syndrome und andere Bewegungsstörungen*. (W. H. Oertel, G. Deuschl, & W. Poewe, Eds.), Stuttgart: Georg Thieme Verlag KG (Vol. 1). Stuttgart: Georg Thieme Verlag. <https://doi.org/10.1055/b-002-41833>
- Olichney, J. M., Murphy, C., Hofstetter, C. R., Foster, K., Hansen, L. A., Thal, L. J., & Katzman, R. (2005). Anosmia is very common in the Lewy body variant of Alzheimer's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, 76(10), 1342–1347. <https://doi.org/10.1136/jnnp.2003.032003>
- Olson, E. J., Boeve, B. F., & Silber, M. H. (2000). Rapid eye movement sleep behaviour disorder: demographic, clinical and laboratory findings in 93 cases. *Brain*, 123(2), 331–339. <https://doi.org/10.1093/brain/123.2.331>
- Postuma, R. B., Montplaisir, J. Y., Pelletier, A., Dauvilliers, Y., Oertel, W., Iranzo, A., ... Wolfson, C. (2012). Environmental risk factors for REM sleep behavior disorder: a multicenter case-control study. *Neurology*, 79(5), 428–434. <https://doi.org/10.1212/WNL.0b013e31825dd383>
- Postuma, Ronald B., Aarsland, D., Barone, P., Burn, D. J., Hawkes, C. H., Oertel, W., & Ziemssen, T. (2012). Identifying prodromal Parkinson's disease: Pre-Motor disorders in Parkinson's disease. *Movement Disorders*, 27(5), 617–626. <https://doi.org/10.1002/mds.24996>
- Postuma, Ronald B., Iranzo, A., Hu, M., Högl, B., Boeve, B. F., Manni, R., ... Pelletier, A. (2019). Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study. *Brain*, 142(3), 744–759. <https://doi.org/10.1093/brain/awz030>
- Pringsheim, T., Jette, N., Frolkis, A., & Steeves, T. D. L. (2014). The prevalence of Parkinson's disease: A systematic review and meta-analysis. *Movement Disorders*, 29(13), 1583–1590. <https://doi.org/10.1002/mds.25945>
- Ritz, B., Lee, P.-C., Lassen, C. F., & Arah, O. A. (2014). Parkinson disease and smoking revisited: Ease of quitting is an early sign of the disease. *Neurology*, 83(16), 1396–1402. <https://doi.org/10.1212/WNL.0000000000000879>
- Rusz, J., Hlavnička, J., Novotný, M., Tykalová, T., Pelletier, A., Montplaisir, J., ... Šonka, K. (2021). Speech Biomarkers in Rapid Eye Movement Sleep Behavior Disorder and Parkinson Disease. *Annals of Neurology*, 90(1), 62–75. <https://doi.org/10.1002/ana.26085>
- Rusz, J., Hlavnička, J., Tykalová, T., Bušková, J., Ulmanová, O., Růžička, E., & Šonka, K. (2016). Quantitative assessment of motor speech abnormalities in idiopathic rapid eye movement sleep behaviour disorder. *Sleep Medicine*, 19, 141–147. <https://doi.org/10.1016/J.SLEEP.2015.07.030>
- Sateia, M. J. (2014). International Classification of Sleep Disorders-Third Edition Highlights and Modifications. *CHEST*, 146(5), 1387–1394. <https://doi.org/10.1378/chest.14-0970>
- Schenck, C. H., Boeve, B. F., & Mahowald, M. W. (2013). Delayed emergence of a

- parkinsonian disorder or dementia in 81% of older men initially diagnosed with idiopathic rapid eye movement sleep behavior disorder: A 16-year update on a previously reported series. *Sleep Medicine*, 14(8), 744–748. <https://doi.org/10.1016/j.sleep.2012.10.009>
- Schenck, C. H., Hurwitz, T. D., & Mahowald, M. W. (1993). REM sleep behaviour disorder: an update on a series of 96 patients and a review of the world literature. *Journal of Sleep Research*, 2(4), 224–231. <https://doi.org/10.1111/j.1365-2869.1993.tb00093.x>
- Schenck, C. H., & Mahowald, M. W. (2002). REM Sleep Behavior Disorder: Clinical, Developmental, and Neuroscience Perspectives 16 Years After its Formal Identification in SLEEP. *Sleep*, 25(2), 120–138. <https://doi.org/10.1093/sleep/25.2.120>
- Schrag, A., Horsfall, L., Walters, K., Noyce, A., & Petersen, I. (2014). Prediagnostic presentations of Parkinson's disease in primary care: a case-control study. *The Lancet Neurology*, 14, 57–64. [https://doi.org/10.1016/S1474-4422\(14\)70287-X](https://doi.org/10.1016/S1474-4422(14)70287-X)
- Sforza, E., Krieger, J., & Petiau, C. (1997). REM sleep behavior disorder: clinical and physiopathological findings. *Sleep Medicine Reviews*, 1(1), 57–69. [https://doi.org/10.1016/S1087-0792\(97\)90006-X](https://doi.org/10.1016/S1087-0792(97)90006-X)
- Shah, M., Muhammed, N., Findley, L. J., & Hawkes, C. H. (2008). Olfactory tests in the diagnosis of essential tremor. *Parkinsonism & Related Disorders*, 14(7), 563–568. <https://doi.org/10.1016/j.parkreldis.2007.12.006>
- Strian, F., & Haslbeck, M. (1986). *Autonome Neuropathie bei Diabetes mellitus*. Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-93310-3>
- Strickland, D., & Bertoni, J. M. (2004). Parkinson's prevalence estimated by a state registry. *Movement Disorders*, 19(3), 318–323. <https://doi.org/10.1002/mds.10619>
- Tsanas, A., Little, M. A., McSharry, P. E., Spielman, J., & Ramig, L. O. (2012). Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease. *IEEE Transactions on Bio-Medical Engineering*, 59(5), 1264–1271. <https://doi.org/10.1109/TBME.2012.2183367>
- Tysnes, O. B., & Storstein, A. (2017). Epidemiology of Parkinson's disease. *Journal of Neural Transmission*, 124(8), 901–905. <https://doi.org/10.1007/s00702-017-1686-y>
- Umsatz mit Smartphones knackt 10-Milliarden-Marke. (2016). Retrieved January 29, 2018, from <https://www.bitkom.org/Presse/Presseinformation/Umsatz-mit-Smartphones-knackt-10-Milliarden-Marke.html>
- Valencia Garcia, S., Brischoux, F., Clément, O., Libourel, P.-A., Arthaud, S., Lazarus, M., ... Fort, P. (2018). Ventromedial medulla inhibitory neuron inactivation induces REM sleep without atonia and REM sleep behavior disorder. *Nature Communications*, 9(1), 504. <https://doi.org/10.1038/s41467-017-02761-0>
- Valencia Garcia, S., Libourel, P.-A., Lazarus, M., Grassi, D., Luppi, P.-H., & Fort, P. (2017). Genetic inactivation of glutamate neurons in the rat sublaterodorsal tegmental nucleus recapitulates REM sleep behaviour disorder. *Brain*, 140(2), 414–428. <https://doi.org/10.1093/brain/aww310>
- Van Den Eeden, S. K. (2003). Incidence of Parkinson's Disease: Variation by Age, Gender, and Race/Ethnicity. *American Journal of Epidemiology*, 157(>11), 1015–1022. <https://doi.org/10.1093/aje/kwg068>
- Visanji, N. P., Brooks, P. L., Hazrati, L.-N., & Lang, A. E. (2013). The prion hypothesis in Parkinson's disease: Braak to the future. *Acta Neuropathologica Communications*, 1(1), 2. <https://doi.org/10.1186/2051-5960-1-2>

- Von Campenhausen, S., Bornschein, B., Wick, R., Bötzel, K., Sampaio, C., Poewe, W., ... Dodel, R. (2005). Prevalence and incidence of Parkinson's disease in Europe. *European Neuropsychopharmacology*, *15*(4), 473–490. <https://doi.org/10.1016/j.euroneuro.2005.04.007>
- Zhang, X., Sun, X., Wang, J., Tang, L., & Xie, A. (2017). Prevalence of rapid eye movement sleep behavior disorder (RBD) in Parkinson's disease: a meta and meta-regression analysis. *Neurological Sciences*, *38*, 163–170. <https://doi.org/10.1007/s10072-016-2744-1>
- Ziemssen, T., Fuchs, G., Greulich, W., Reichmann, H., Schwarz, M., & Herting, B. (2011). Treatment of dysautonomia in extrapyramidal disorders. *Journal of Neurology*, *258*(2), 339–345. <https://doi.org/10.1007/s00415-011-5946-8>

10 Anhang

10.1 Informationsblatt zur Messungsdurchführung

Philipps-Universität Marburg

Fachbereich Medizin



Klinik für Neurologie

Prof. Dr. Dr. h.c. W. H. Oertel
Hartie Senior-Forschungsprofessor

Informationsblatt zur Durchführung der Smartphone-basierten Messungen zu Hause

Studie: „Untersuchung der Stimmcharakteristik, Körperbewegung und Geschicklichkeit bei Patienten mit REM-Schlaf-Verhaltensstörung (RBD), Parkinson-Krankheit (PK), Multi-System-Atrophie (MSA), Supranuklearer Blickparese (PSP) und gesunden Kontrollpersonen“.

In Kooperation mit der Oxford Universität, Vereinigtes Königreich

Sehr geehrter Studienteilnehmer,
sehr geehrte Studienteilnehmerin,

vielen Dank, dass Sie sich bereit erklärt haben uns bei unserer Smartphone-basierten Untersuchung zur Stimmcharakteristik, Körperbewegung und Geschicklichkeit zu unterstützen.

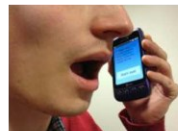
Auf diesem Informationsblatt erhalten Sie im **ersten Teil A** Instruktionen zum Durchführen der sieben, kurzen „Mini-Tests“, die Sie bereits mit uns in der Klinik durchgeführt haben. Im **zweiten Teil B** finden Sie Informationen zur „Passiven Aufzeichnung“, d.h. die Aufzeichnungen ihrer Körperbewegungen, die Sie bitte in der gleichen Messwoche zu Hause einschalten während Sie wach sind.

A. Aktive Mini-Tests

Im Folgenden finden Sie ausführliche Instruktionen zum Durchführen der „aktiven Mini-Tests“, die Sie bitte möglichst in regelmäßigen Abständen vier Mal täglich eine Woche lang durchführen.

(1) Stimmtest

Bitte suchen sie einen ruhigen, ungestörten Ort auf, der möglichst frei von technischen Nebengeräuschen wie z.B. durch eine Klimaanlage, Pumpenanlagen oder ähnliches ist. Die Instruktion ist die selbe wie Sie sie erhalten haben als Sie den Stimmtest bei uns in der Klinik durchgeführt haben: „Drücken Sie den ‚Start Test‘-Button, halten Sie das Handy an Ihr Ohr als würden Sie ein Telefonat führen wollen, atmen Sie tief ein und sagen Sie daraufhin ‚aaah‘ so gleichmäßig und lange Sie können.“ Die Aufzeichnung beginnt sofort nachdem Sie den ‚Start Test‘-Button gedrückt haben und hält 20 Sekunden lang an. Sagen Sie so lange ‚aaah‘ bis Sie außer Atem sind oder auf natürliche Weise aufhören. Eine Aufzeichnung von mehr als sechs Sekunden ist ideal – Sie müssen nicht die gesamten 20 Sekunden der Aufzeichnung durchhalten!



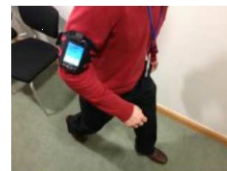


(2) Gleichgewichtstest

Nun suchen Sie bitte einen Platz auf, an dem Sie freihändig und ohne Hilfe stehen können. Drücken Sie den ‚Start-Test‘-Button. Legen Sie das Handy nun bitte in Ihre Hosentasche und stellen sich dann aufrecht hin. Das Handy wird nun einmal vibrieren – bleiben Sie derweil bitte aufrecht stehen. Wenn das Handy ein zweites Mal vibriert hat, können Sie es wieder aus der Hosentasche nehmen. Sollten Sie keine Hosentaschen tragen, benutzen Sie einfach den Armgurt, den wir Ihnen mitgegeben haben.

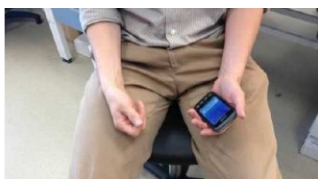
(3) Gangtest

Jetzt suchen Sie bitte einen Ort auf, wo Sie ungehindert ca. 18 Meter gehen können und drücken wieder den ‚Start-Test‘-Button. Stellen Sie sich nun aufrecht hin und platzieren Sie das Handy in Ihrer Hosentasche. Wenn der Buzzer vibriert, gehen Sie bitte die 18 Meter geradeaus, drehen dann um und gehen den gleichen Weg wieder zurück. Auch hier können Sie wieder den Armgurt verwenden, wenn Sie gerade keine Hosentaschen tragen.



(4) Ruhetremortest

Setzen Sie sich bitte aufrecht auf einen Stuhl und halten Sie das Handy in der Hand, die am meisten vom Tremor betroffen ist. Sollten Sie nicht unter einem Tremor leiden, halten Sie das Smartphone einfach in eine Hand Ihrer Wahl. Legen Sie dann beide Hände leicht auf Ihre Oberschenkel, so dass diese leicht auf Ihrem Schoß ruhen können. Drücken Sie den ‚Start-Test‘-Button, wenn Sie soweit sind und schließen danach die Augen. Zählen Sie bitte von 100 rückwärts (100 – 99 – 98 – 97 – und so weiter) während Sie weiterhin das Handy in Ihrer Hand halten. Das Telefon vibriert einmal, wenn die Aufzeichnung beginnt und vibriert ein zweites Mal, wenn diese nach ca. 45 Sekunden beendet ist.





(5) Haltetremortest

Auch bei diesem Test setzen Sie sich bitte aufrecht hin und halten das Telefon in der Hand, die am meisten vom Tremor betroffen ist. Sollten Sie nicht unter einem Tremor leiden, halten Sie das Smartphone wieder einfach in eine Hand Ihrer Wahl. Drücken Sie den ‚Start-Test‘-Button, wenn Sie soweit sind und halten Sie die Hand mit dem Telefon danach gerade gestreckt vor sich. Das Handy vibriert auch hier einmal, wenn die Aufzeichnung beginnt, diese hält für 45 Sekunden an und wird durch ein zweites Vibrieren beendet. Bitte machen Sie sich keine Sorgen, wenn Sie Ihre Hand nicht die ganzen 45 Sekunden gestreckt vor sich halten können; versuchen Sie es einfach so gut es geht.

(6) Reaktionstest

Legen Sie das Handy auf eine flache Oberfläche und drücken Sie den ‚Start-Test‘-Button, wenn Sie soweit sind. Immer wenn am unteren Rand des Bildschirms ein (rechteckiges) Feld erscheint, tippen Sie so schnell wie möglich mit dem Finger darauf und halten Ihren Finger solange auf dem Feld gedrückt bis es wieder vom Bildschirm verschwindet. Erscheint das Feld wieder auf dem Bildschirm, tippen Sie mit Ihrem Finger wieder darauf, halten es gedrückt und nehmen den Finger schnellstmöglich vom Bildschirm, sobald das Feld wieder vom Bildschirm verschwindet. Wiederholen Sie dies solange bis die App die Messung automatisch beendet.

(7) Geschicklichkeitstest

Auch hier legen Sie das Handy bitte auf eine flache Oberfläche und drücken den ‚Start-Test‘-Button, wenn Sie soweit sind. Tippen Sie nun abwechselnd mit dem Zeige- und Mittelfinger einer Hand im gleichmäßigen Rhythmus auf die beiden Felder unten auf dem Bildschirm (als wollten Sie mit zwei Fingern Klavier spielen). Hierfür benutzen Sie bitte Ihre dominante Hand (z.B. die rechte Hand, wenn Sie Rechtshänder sind). Sollten Sie beidhändig sein, benutzen Sie bitte die Hand, mit der Sie normalerweise Schreiben.



B. Passives Monitoring

Neben den Smartphone-basierten Messungen mittels der sieben kurzen „Mini-Tests“, möchten wir Sie bitten innerhalb der gleichen Messwoche die „passive Aufzeichnung“ so oft wie möglich einzuschalten. Nach dem Informationsblatt_Instruktionen_zuhause_Version_1_Stand_14.09.2015

Philipps-Universität Marburg
Fachbereich Medizin



Klinik für Neurologie
Prof. Dr. Dr. h.c. W. H. Oertel
Hartje Senior-Forschungsprofessor

Öffnen der App tippen Sie dafür bitte auf das Feld „Monitoring einschalten“. Danach beginnt die Aufzeichnung. Bitte tragen Sie das Handy während der Aufzeichnungen möglichst am Körper – entweder in Ihrer Hosen- bzw. Jackentasche oder Sie benutzen den Handy-Tragegurt, den wir Ihnen zur Verfügung gestellt haben. Nach Einschalten der Aufzeichnung zeigt Ihnen die App an, dass die Aufzeichnung gerade aktiv ist und wann diese gestartet wurde. Möchten Sie die Aufzeichnung beenden, tippen Sie einfach auf das Feld „Monitoring ausschalten“. Die kontinuierliche Aufzeichnung läuft „im Hintergrund“, d.h. Sie können Ihr Smartphone währenddessen wie gewohnt nutzen.

Die erfassten Daten der „Mini-Tests“ und des „passiven Monitorings“ werden automatisch verschlüsselt via Internet von Ihrem Smartphone auf den projektbezogenen Server übermittelt, sobald sich Ihr Gerät in der Nähe eines verfügbaren WLAN-Zugangs befindet. Ist Ihr Handy mit dem WLAN verbunden, wird die App alle 15 Minuten automatisch einen Upload durchführen und die bis dahin erhobenen Daten auf den gesicherten Server übermitteln. Sie können die Übermittlung der Daten aber auch manuell durchführen, indem Sie im Haupt-Menü der Smartphone-App auf den Menüpunkt „Hochladen“ tippen. Nach ungefähr 30 Sekunden erscheint dann auf dem Bildschirm eine Benachrichtigung, dass die Übermittlung der Daten erfolgreich war.

Nach den 7 Messtagen bei Ihnen Zuhause haben wir genügend Daten.

Sollten Sie Fragen haben oder sich während der Messwoche Probleme ergeben, können Sie gerne Frau Dipl.-Psych. Karén Wilhelm unter der Telefonnummer 06421-58-63798 erreichen oder Sie schreiben eine E-Mail an karen.wilhelm@staff.uni-marburg.de.

Wir möchten uns bereits jetzt für Ihr Engagement und Ihre Unterstützung für dieses Forschungsprojekt bedanken.

Mit freundlichen Grüßen,
 Ihr Team der Klinik für Neurologie, Marburg

10.2 Informationen und Aufklärung für Studienteilnehmer

Philipps-Universität Marburg

Fachbereich Medizin



Klinik für Neurologie

Prof. Dr. Dr. h.c. W. H. Oertel
Hertie Senior-Forschungsprofessor

Informationen für die Studienteilnehmer

Smartphone-basierte Untersuchung zur Stimmcharakteristik, Körperbewegung und Geschicklichkeit bei Patienten mit REM-Schlaf-Verhaltensstörung (RBD), Parkinson-Krankheit (PK), Multi-System-Atrophie (MSA), progressiver supranuklearer Blickparese (PSP) und gesunden Kontrollpersonen.

Verantwortliche Versuchsleiter

Prof. Dr. med. Dr. h.c. Wolfgang Oertel
Fachbereich Humanmedizin
Philipps-Universität Marburg
Klinik für Neurologie
Baldingerstraße
35043 Marburg

PD Dr. med. Christoph Best
Fachbereich Humanmedizin
Philipps-Universität Marburg
Klinik für Neurologie
Baldingerstraße
35043 Marburg

Durchführender Versuchsleiter: Moritz Beilharz

Allgemeine Informationen:

Wir bitten Sie um Teilnahme an einer Studie. Dieses Formular dient dem Zweck, Ihnen alle Informationen zu geben, die erforderlich sind, damit Sie entscheiden können, ob Sie bereit sind an der Studie teilzunehmen oder nicht. Bitte lesen Sie sich das Formular aufmerksam durch. Sollten für Sie, nachdem Sie alles gelesen haben, nicht alle Fragen beantwortet sein, dann zögern Sie bitte nicht, restliche Unklarheiten anzusprechen und mit dem/der Versuchsleiter/in auszuräumen. Wenn alle Ihre Fragen beantwortet sind, können Sie sich entscheiden, ob Sie an der Studie mitwirken möchten oder nicht.

Die Teilnahme an der Studie ist freiwillig. Sie können jederzeit und ohne Angaben von Gründen Ihre Einwilligung zur Studienteilnahme widerrufen, ohne dass Ihnen daraus Nachteile entstehen.

Zweck der Studie:

Ziel dieser Studie ist es, mit Smartphone-gesteuerten „Mini“-Tests Eigenschaften der Sprache, Körperbewegung sowie der Geschicklichkeit der Hände bei Patienten mit REM-Schlafverhaltensstörung (RBD) digital aufzuzeichnen und somit objektive Daten zu erhalten, um die Symptomatik zuverlässig messbar und vergleichbar zu machen. Dazu sollen die Daten der RBD-Patienten erfasst und mit den Ergebnissen von gesunden Kontrollpersonen sowie Patienten mit bekannter Parkinson-Krankheit (PK), Multi-System-Atrophie (MSA) und Patienten mit einer



supranuklearen Blickparese (PSP) verglichen werden. Sofern die Ergebnisse systematische Unterschiede aufzeigen, könnte man mit dieser Messmethode eine mögliche diagnostische Früherkennung für neurodegenerative Erkrankungen sowie eine schnelle, unkomplizierte und zuverlässige Methode zur Symptombeobachtung etablieren.

Ablauf der Studie:

Im Rahmen dieser Untersuchung benutzen Sie ein von uns gestelltes Leih-Smartphone, welches eine speziell für diese Untersuchung entwickelte Smartphone-App (ein Programm auf einem Mobilfunkgerät) beinhaltet.

Diese App umfasst insgesamt sieben kurze und leicht durchführbare „Mini-Tests“:

1) *Stimmtest* 2) *Gleichgewichtstest*, 3) *Gangtest*, 4) *Ruhetremortest*, 5) *Halbtremortest*, 6) *Reaktionstest* und einen 7) *Geschicklichkeitstest*.

Darüber hinaus beinhaltet die App eine „passive Aufzeichnung“. Damit können bestimmte alltägliche Bewegungsmuster und Aktivitätsstufen aufgezeichnet und später ausgewertet werden. Diese Aufzeichnung läuft im Hintergrund ab. Ihre alltäglichen Aktivitäten oder die Nutzung des Telefons werden davon nicht beeinflusst. Auch der Inhalt Ihrer Telefongespräche wird in keiner Weise aufgezeichnet.

Für die Studie bitten wir Sie A) zum einen die sieben „Mini-Tests“ der App während eines Kontrolltermins bei Ihrem behandelnden Arzt hier in der Klinik durchzuführen.

Zusätzlich benutzen Sie die Smartphone-App B) darauffolgend eine Woche lang zu Hause, um damit eigenständig vier Mal täglich in regelmäßigen Abständen eine Messung mit den sieben „Mini-Tests“ durchzuführen.

Dies tun Sie zum Beispiel wie folgt:

8 Uhr: 7 Minitests
 11 Uhr: 7 Minitests
 15 Uhr: 7 Minitests
 19 Uhr: 7 Minitests

Selbstverständlich können Sie entsprechend Ihrem individuellen Tagesablauf auch andere Zeiten wählen. Wir bitten Sie dann, diese einmal von Ihnen gewählten und festgelegten Zeitpunkte an allen 7 Tagen für die Durchführungen der Messungen in gleicher Weise einzuhalten. Falls es aber nicht möglich ist, die Minitests immer zu den gleichen Zeiten durchzuführen, können Sie auch selbst an jedem der 7 Tage flexibel entscheiden, wann Sie die 4-maligen Testungen am Tage durchführen.

Ein Messdurchgang mit allen sieben „Mini-Tests“ dauert gewöhnlich nicht länger als sieben Minuten. Die eigentliche Messzeit beträgt drei Minuten.





Darüber hinaus aktivieren Sie so oft wie möglich in dieser Untersuchungswoche zuhause eine unauffällige „passive Aufzeichnung“ während Ihrer Wachstunden und tragen dafür das Smartphone kontinuierlich bei sich am Körper. Dafür stecken Sie das Handy entweder in Ihre Jacken- oder Hosentasche oder Sie benutzen einen Handy-Tragegurt, den wir Ihnen zur Verfügung stellen.

Erfasste Daten:

Die Smartphone-App benutzt Sensoren, die bereits in jedem handelsüblichen Smartphone werkmäßig enthalten sind. Dazu gehören Bewegungssensoren, Annäherungssensoren, das Gyroskop (Kreiselsensor) und das Mikrofon. Das Mikrofon wird ausschließlich während des Stimmtests für die Stimmaufzeichnung aktiviert. Während des Gleichgewichts-, Gang-, Ruhetremor- und Haltetremortests werden die Bewegungssensoren aktiviert, um Bewegungsamplituden und Beschleunigungen der Bewegung zu erfassen. Der Reaktions- und Geschicklichkeitstest erfasst hingegen Reaktionszeiten Ihrer Berührungen auf dem Touch-Bildschirm. Neben den Bewegungssensoren wird bei der passiven Aufzeichnung auch das GPS-System (Globales Positionsbestimmungssystem) aktiviert, um neben Bewegungsmustern und Aktivitätsniveaus auch Gangbeschleunigungen bergauf oder bergab zu erfassen.

Datenschutz und Datenspeicherung:

Alle Daten werden unter Einhaltung des Datenschutzgesetzes erhoben, gespeichert und verarbeitet. Die Daten, die von der Smartphone-App elektronisch erfasst werden, werden verschlüsselt auf einem projektbezogenen Server übermittelt, gespeichert und analysiert.

Personenbezogene Daten werden von der Smartphone-App nicht erfasst. Zugriff auf die Daten haben nur unmittelbar mit der Studie befasste Personen sowie autorisierte Personen der zuständigen Ethikkommission zu Prüfzwecken unter Wahrung der Vertraulichkeit.

Wir möchten Sie an dieser Stelle darauf aufmerksam machen, dass aus soeben genannten datenschutzrechtlichen Gründen die Nutzung des eigenen Smartphones nicht möglich ist.

Die Speicherung und Auswertung der Messdaten erfolgt in elektronischer Form. Hierbei wird Ihren Messdaten eine Nummer zugeordnet („Pseudonymisierung“). Die Zuordnung zwischen dieser Nummer („Pseudonym“) und personenbezogenen Daten erfolgt ausschließlich in Papierform in einfacher Ausfertigung; die zugehörigen Dokumente werden in verschlossenen Schränken, zu denen nur die Versuchsleiter Zugang haben, verwahrt. Daten auf elektronischen Speichermedien werden durch Passwortschutz vor unerlaubtem Zugriff geschützt; die Speichermedien selbst werden in verschlossenen, nur dem oben genannten Personenkreis zugänglichen, Räumen aufbewahrt. Gemäß gesetzlichen Regelungen werden alle Daten zehn Jahre aufbewahrt und anschließend vernichtet. Alle Untersucher/innen sowie die anderen genannten mit der Verarbeitung der Daten betrauten Personen unterliegen der Verschwiegenheitsverpflichtung nach §40 Bundesdatenschutzgesetz.

Mögliche Risiken, Stress oder Unannehmlichkeiten:

Mit dieser Studie sind keine uns bekannten Risiken verbunden. Die Testungen sind unkompliziert, nicht schmerzhaft und werden in der Regel auch nicht als unangenehm empfunden. Sie könnten es eventuell als etwas stressig empfinden, in der Messwoche zu Hause die vier täglichen, regelmäßigen Messungen mit den „Mini-Tests“ in Ihrem Alltag zu integrieren. Falls erforderlich, können Sie aber flexibel selbst entscheiden, wann und wo Sie diese Messungen durchführen. Weiterhin könnten Sie es als unangenehm empfinden während der „passiven Aufzeichnung“ das Smartphone kontinuierlich am Körper zu tragen. Aber auch hier

Philipps-Universität Marburg
Fachbereich Medizin



Klinik für Neurologie
Prof. Dr. Dr. h.c. W. H. Oertel
Hertie Senior-Forschungsprofessor

entscheiden Sie zu jeder Zeit selbst, wie oft und wie lange Sie die kontinuierliche Aufzeichnung aktivieren. Zudem haben Sie während der Aufzeichnung natürlich weiterhin die Möglichkeit Ihr Smartphone wie gewohnt zu benutzen.

Es entstehen Ihnen keinerlei Nachteile, wenn Sie sich dazu entschließen, eine Messung nicht durchzuführen oder die „passive Aufzeichnung“ auszuschalten.

Nutzen des Versuches:

Für Sie wird aus der Versuchsdurchführung kein weiterer Nutzen resultieren. Wir erhoffen uns aber aus den Ergebnissen ein objektives, diagnostisches Hilfsmittel zu entwickeln, welches zukünftig möglicherweise die Frühdiagnostik bei neurodegenerativen Erkrankungen erleichtern und die behandlungsbegleitende Verlaufskontrolle von motorischen und nicht motorischen Symptomen bei der Parkinson-Erkrankung verbessern kann.

Deklaration von Helsinki:

Der Versuch steht in Einklang mit den in der Deklaration von Helsinki niedergelegten ethischen Standards für die Forschung an Menschen.

Im Falle von Rückfragen, können Sie jederzeit mit dem verantwortlichen Versuchsleiter Kontakt aufnehmen. Sie erreichen uns unter o.g. Adresse, unter Telefon 06421-58 63798 oder via Email an karen.wilhelm@staff.uni-marburg.de!

Bitte noch auf der nächsten Seite die Einverständniserklärung unterschreiben —>

Philipps-Universität Marburg
Fachbereich Medizin



Klinik für Neurologie
 Prof. Dr. Dr. h.c. W. H. Oertel
 Hartig Senior-Forschungsprofessor

Einverständniserklärung

Smartphone-basierte Untersuchung zur Stimmcharakteristik, Körperbewegung und Geschicklichkeit bei Patienten mit REM-Schlaf-Verhaltensstörung (RBD), Parkinson-Krankheit (PK), Multi-System-Atrophie (MSA), progressiver supranukleare Blickparese (PSP) und gesunden Kontrollpersonen.

Ich habe die Probandeninformation über Ziel und Ablauf der Untersuchung sowie studienbedingte Erfordernisse und mögliche Nebenwirkungen erhalten, gründlich durchgelesen und verstanden. Ich hatte ausreichend Gelegenheit, mich bei dem/der Versuchsleiter/Versuchsleiterin über den Untersuchungshergang zu informieren, sowie auftretende Fragen zu stellen. Diese wurden mir von dem/der Versuchsleiter/Versuchsleiterin verständlich beantwortet. Eine Kopie der Probandeninformation habe ich erhalten. Ich hatte ausreichend Zeit, mich für oder gegen eine Teilnahme an dieser Studie zu entscheiden. Mit meiner Unterschrift erkläre ich, dass ich das Vorhaben und die Information verstanden habe und freiwillig an der Studie teilnehme. Ich habe verstanden, dass ich jederzeit ohne Angabe von Gründen aus der Studie ausscheiden kann, ohne dass mir persönliche Nachteile entstehen. Auch der Versuchsleiter kann die Studie jederzeit beenden. Mir ist bekannt, dass diese Studie in erster Linie der Wissenserweiterung dient und gegebenenfalls keinen persönlichen Vorteil für mich bringen kann. Ich erkläre mich damit einverstanden, dass meine Daten unter Einhaltung des Datenschutzgesetzes erhoben, gespeichert und verarbeitet werden. Ich bin darüber informiert, dass alle Untersucher/innen der Verschwiegenheitsverpflichtung nach §40 Bundesdatenschutzgesetz unterliegen und die Speicherung und Auswertung meiner studienbezogenen Daten nach gesetzlichen Bestimmungen in anonymisierter oder pseudonymisierter Form erfolgt.

Versuchsperson:

Datum	Name in Blockschrift	Unterschrift
-------	----------------------	--------------

Versuchsleiter/Versuchsleiterin:

Datum	Name in Blockschrift	Unterschrift
-------	----------------------	--------------

10.3 Ethikantrag



**Kommission für Ethik in der ärztlichen Forschung
Fachbereich Humanmedizin der Philipps-Universität Marburg
Antragsformular**

1. Synopsis:

Titel: „Smartphone-basierte Untersuchung zur Stimmcharakteristik, Körperbewegung und Geschicklichkeit bei Patienten mit REM-Schlaf-Verhaltensstörung (RBD), Parkinson-Krankheit (PK) und Multi-System-Atrophie (MSA), bei einer Krankheitskontrolle (progressive supranukleäre Blickparese (PSP)) und bei gesunden Kontrollpersonen“

Die wissenschaftliche Fragestellung lautet:

Methodik: Wir erfassen objektive, standardisiert durchgeführte Bewegungs- und Sprech-Signale (Artikulation, Modulation) mittels eines ambulant getragenen, handelsüblichen Smartphones (Accelerometer für Bewegung, Mikrofon für die Sprache). Diese Methode erfasst Störungen der Motorik (Innervation von Extremitäten- und Rumpfmuskeln, Sprech-Muskeln).

Frage: Erlaubt diese Methode, motorische Störungen bei Patienten mit einer REM-Schlafverhaltensstörung, die ein sehr hohes Risiko (>85%) haben, später an einer Parkinson-Krankheit oder Multi-System-Atrophie zu erkranken, früher zu erkennen als der erfahrene Neurologe, der mit Hilfe des „klinischen Blickes“ und einer Beurteilungsskala motorische Symptome erfasst?

Projektleiter:

Prof. Dr. med. Dr. h.c. Wolfgang Hermann Oertel

Hertie Senior Forschungsprofessor

Klinik für Neurologie, FB Medizin, Philipps-Universität Marburg

Adresse: UKGM, Klinik für Neurologie, Baldingerstraße, 35043 Marburg

Telefon: 06421 - 58 - 6 3731/ 5215/ 5217

Fax: 06421 - 58 - 6 8955

E-Mail: oertelw@med.uni-marburg.de

2. Studientyp/Studiendesign:

Experimentelle Querschnitt-Studie

(diese Studie soll später in eine Längsschnitts-Studie mit insgesamt 6 Messzeitpunkten in 9-monatigen Abständen übergehen – hierzu wird ein gesonderter Antrag an die Ethikkommission gestellt – siehe 3B)



3. Zeitplan:

A) Querschnitt

18 Monate (September 2015 bis Februar 2017)

B) Längsschnitt (mögliches Folgeprojekt; nicht Teil dieses Antrages)

Es ist geplant die Querschnitt-Studie später in eine Längs-Schnittstudie mit insgesamt 6 weiteren Messzeitpunkten in 9-monatigen Messabständen übergehen zu lassen.

4. Patientengruppe:

Patienten mit einer REM-Schlaf-Verhaltensstörung (RBD)

Patienten mit einer Parkinson-Krankheit (PK)

Patienten mit einer Multi-System-Atrophie (MSA)

Patienten mit einer Progressiven Supranuklearen Blickparese (PSP) (Krankheitskontrolle)

Gesunde Kontrollpersonen



A Formales

1. Titel der Studie

Smartphone-basierte Untersuchung zur Stimmcharakteristik, Körperbewegung und Geschicklichkeit bei Patienten mit REM-Schlaf-Verhaltensstörung (RBD), Parkinson-Krankheit (PK), Multi-System-Atrophie (MSA), bei einer Patientengruppe als Krankheitskontrolle (progressiver supranuklearer Blickparese (PSP)) und gesunden Kontrollpersonen

1. Verantwortlicher Studienleiter in Marburg

Prof. Dr. med. Dr. h.c. Wolfgang Oertel, Hertie Senior Forschungsprofessor,
Klinik für Neurologie, FB Humanmedizin, Philipps-Universität Marburg
Adresse: UKGM, Klinik für Neurologie, Baldingerstraße, 35043 Marburg
Telefon: 06421 – 58 – 6 3731
E-Mail: oertelw@med.uni-marburg.de

Vertreter:

PD Dr. Christoph Best, Oberarzt und wissenschaftlicher Mitarbeiter,
Klinik für Neurologie, FB Humanmedizin, Philipps-Universität Marburg
Adresse: UKGM, Klinik für Neurologie, Baldingerstraße, 35043 Marburg
Telefon: 06421 – 58 – 6 1877
E-Mail: bestc@uni-marburg.de

2. Sonstige Beteiligte

Oxford, Vereinigtes Königreich

Dr. Max Little, Assistant Professor, Aston University, UK
Adresse: Aston Triangle, Birmingham, B4 7ET, Vereinigtes Königreich
Telefon: 0044 – 121 – 204 – 5327
E-Mail: max.little@aston.ac.uk



Dr. Siddharth Arora, post-doctoral Research Fellow, Somerville College, University of Oxford,
UK

Adresse: Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road,
Oxford OX2 GG6, Vereinigtes Königreich

Telefon: 0044 – 186 – 561 – 5126

E-Mail: siddharth.arora@some.ox.ac.uk

Dr. Michele Hu, Associate Professor and Consultant Neurologist, Nuffield Department of
Clinical Neurosciences, Oxford Parkinson's Disease Centre, University of Oxford, UK

Adresse: Level 6, West Wing, John Radcliffe Hospital, Oxford OX3 9DU,
Vereinigtes Königreich

Telefon: 0044 – 186 – 523 – 1295

E-Mail: michele.hu@ndcn.ox.ac.uk

Dr. Fahd Baig, Clinical Research Fellow, Oxford Parkinson's Disease Centre, Nuffield
Department of Clinical Neurosciences, University of Oxford, UK

Adresse: Level 6, West Wing, John Radcliffe Hospital, Oxford OX3 9DU,
Vereinigtes Königreich

Telefon: 0044 – 777 – 376 – 6013

E-Mail: fahd.baig@ndcn.ox.ac.uk

Klinik Für Neurologie, Marburg

Dr. David Vadasz, Assistenzarzt und wissenschaftlicher Mitarbeiter, Klinik für Neurologie,
FB Humanmedizin, Philipps-Universität Marburg

Dipl.-Psych. Karén Wilhelm, wissenschaftliche Mitarbeiterin, Klinik für Neurologie,
FB Humanmedizin, Philipps-Universität Marburg

3. Multizentrische Studie

Bei dem beantragten Projekt handelt es sich um eine multizentrische Studie. Neben der
beantragten Untersuchung in Marburg, werden gleiche Untersuchungen in Oxford,
Vereinigtes Königreich, durchgeführt.



4. Auftraggeber

Es gibt keinen Auftraggeber und/oder industriellen Sponsor.

5. Finanzierung

Die Finanzierung der beantragten Untersuchung am Standort Marburg erfolgt über die Ausstattungsmittel der Hertie-Senior-Forschungsprofessur (Prof. Dr. Dr. Wolfgang H. Oertel).

6. Prüfvertrag / Prüfarzt

Nein.

7. Wurde die gleiche Studie schon einer Ethik-Kommission vorgelegt?

Das Projektvorhaben wurde bereits der Ethikkommission in Oxford, Vereinigtes Königreich, vorgelegt. Ein positives Ethikvotum liegt dort unter der Referenznummer 10/H0505/71 bereits vor (eine Kopie des positiven Ethikvotums vom 22. November 2013 der dort zuständigen Ethikkommission wurde dem Antrag im Anhang beigelegt).

B Projektbeschreibung

8. Studienvorhaben: „Smartphone-basierte Untersuchung zur Stimmcharakteristik, Körperbewegung und Geschicklichkeit bei Patienten mit REM-Schlaf-Verhaltensstörung (RBD) – einer prodromalen Phase der PK und MSA -, mit Parkinson-Krankheit (PK) und Multi-System-Atrophie (MSA) sowie bei einer Krankheitskontrolle (progressive supranukleare Blickparese (PSP)) und bei gesunden Kontrollpersonen“

Stand des Wissens / Gelöste und ungelöste Probleme in der Literatur

Die Parkinson-Erkrankung (*Parkinson's disease*, PK) ist mit einer europaweiten Prävalenzrate von zwischen 100-200/100.000 Einwohnern eine der häufigsten neurodegenerativen Erkrankungen (7). Die Prävalenzrate ist deutlich altersabhängig und im Zuge des demografischen Wandels wird geschätzt, dass sich diese bei Personen über 50 Jahren von 2010 bis 2030 weltweit verdoppeln wird (6). Die PK wird klinisch durch motorische Symptome wie die Verlangsamung von Bewegungsabläufen (Akinese), Rigor, Tremor in Ruhe



und - im fortgeschrittenen Stadium - durch eine eingeschränkte Stabilität der Körperhaltung charakterisiert. Es wird geschätzt, dass die ersten Anzeichen derartiger motorischer Symptome erst dann auftreten, wenn bereits die Hälfte der dopaminergen Neurone in der Substantia Nigra und ca. 70 % des dopaminergen Gewebes im Putamen verloren gegangen sind (8, 9). Aufgrund der Arbeiten von Braak et al. 2003 wird mittlerweile davon ausgegangen, dass der neurologisch definierten, d.h. motorischen Phase der Parkinson-Krankheit eine langjährige prodromale (nicht-motorische) Phase vorausgeht. Diese wird vereinfacht in drei Phasen eingeteilt: Phase 1 ist durch Hyposmie und autonome Dysfunktionen (12, 13) z.B. des gastrointestinalen Traktes gekennzeichnet, Phase 2 werden die REM-Schlafverhaltensstörung (*rapid eye movement sleep behaviour disorder*, RBD (5)) und affektive Störungen wie Depressionen (10, 11) zugeordnet, während in Phase 3 der nigrostriatale dopaminerge Trakt beginnt zu degenerieren und erste sehr milde „motorische“ Störungen beobachtet werden können.

Die RBD ist eine Schlafstörung, die durch fehlende Muskeltonie in der Traumschlafphase gekennzeichnet ist. Dadurch kommt es während des REM-Schlafes zu oftmals komplexen motorischen Bewegungsabläufen, meist in Form von einem Ausagieren aggressiver Traumhalte. Es leiden schätzungsweise 0,5 % der allgemeinen Bevölkerung an dieser Schlafstörung (1). Im weiteren Krankheitsverlauf konvergiert die RBD in mehr als 85 % der Fälle in eine neurodegenerative Erkrankung vom Typ der Alpha-Synucleinopathien. So entwickeln Patienten mit einer idiopathischen RBD innerhalb von 10 bis 20 Jahren entweder eine PK oder in selteneren Fällen eine Multi-System-Atrophie (*multiple system atrophy*, MSA) oder eine Demenz mit Lewy-Körpern (*dementia with Lewy bodies*, DLB) (2, 3, 4). Die RBD stellt somit den bisher spezifischsten Indikator für die Entstehung einer späteren neurodegenerativen Erkrankung dar.

Sprechstörungen bei Parkinsonpatienten, gemessen mittels aufwendiger Tonbandaufzeichnungen, zeigen sich bereits bis zu fünf Jahre vor der Diagnosestellung einer PK (22). Und auch erste leichte, motorische Besonderheiten werden bei RBD-Patienten bereits bis zu 4,5 Jahren vor der offiziellen Diagnosestellung „Parkinson-Syndrom“ erkennbar (23).

In der Diagnose der PK sowie ihres prodromalen Stadiums RBD ist es daher von großer Bedeutung, motorische wie auch nicht motorische Auffälligkeiten möglichst frühzeitig,



vollständig, objektiv, valide und kostengünstig erfassbar zu machen, 1) um den weiteren Krankheitsverlauf, das heißt den Übergang von der prodromalen in die manifeste Phase der Parkinson-Krankheit präzise dokumentieren und damit verfolgen zu können, 2) um spezifische in Entwicklung befindliche oder noch zu entwickelnde neuroprotektive Therapien kontrolliert prüfen zu können und 3) nach der erfolgten Zulassung frühzeitig, d.h. im prodromalen Stadium der PK einsetzen zu können.

Studienbegründung

Bislang beruhen die gültigen klinischen Diagnosekriterien auf dem Nachweis der oben erwähnten motorischen Kardinalsymptome der PK, die durch eine Kombination von objektiven klinischen (das heißt teilweise doch subjektiven, da auf großer Erfahrung beruhenden) Untersuchungsbefunden und einem subjektiven klinischen Urteil (Blickdiagnose) der Fachärzte erhoben werden. Zur Erfassung des Schweregrades und zur Einschätzung der Therapieeffekte hat sich – vor allem im wissenschaftlichen Bereich – die „Unified Parkinson Disease Rating Scale“ (UPDRS (17)) durchgesetzt. Objektive, symptom-spezifische diagnostische Hilfsmittel liegen derzeit noch nicht ausreichend vor. Vor allem würde die Diagnostik mittels Smartphone-basierter Tests den großen Vorteil bieten, dass diese Tests symptom-spezifische, objektive und zuverlässige skalierungsfähige Maße für die therapiebegleitenden Verlaufskontrollen liefern und leicht durch den Patienten zuhause durchführbar sind.

Vorangegangene Studien konnten zeigen, dass es möglich ist, mit digitalen Aufzeichnungen der Stimme und der Körperbewegung durch handelsübliche Smartphones sehr genau (mit 98 - 99 % Sensitivität und Spezifität und einer 5 %-Fehlerquote gemessen zur UPDRS) zwischen PK-Patienten und gesunden Kontrollpersonen zu unterscheiden (14, 15, 16). Weitere Studien, die Verhaltensweisen wie z.B. Körperbewegungen und Mobilität im Alltag mit Hilfe von Smartphones erfasst hatten, konnten die oben erwähnten Ergebnisse unabhängig replizieren (18, 19). Darüber hinaus ist es durch die Anwendung speziell entwickelter, mathematischer Logarithmen gelungen, den Schweregrad parkinson-typischer Symptome objektiv quantifizierbar zu machen (21).



Welche Studien fehlen?

- 1) Smartphone-basierte publizierte Untersuchungen zur Sprechanalyse bei RBD-Patienten als Prodromalphase der PK liegen nach unserem Wissen nicht vor.
- 2) Studien speziell mit Smartphone-basierten Daten zur Unterscheidung zwischen der PK und den atypischen Parkinson-Syndromen (MSA, PSP) liegen nach unserem Wissen nicht vor. Eine aktuelle Untersuchung konnte zeigen, dass man mit Hilfe von Sprechaufzeichnungen per Tonband signifikant zwischen Patienten mit PK, MSA und PSP unterscheiden kann (20).
- 3) Es gibt keine kombinierte Untersuchung der Stimmcharakteristik, Körperbewegungen und Geschicklichkeit bei Patienten mit einer RBD, die sich – wie bereits oben erwähnt– als Prodromalphase von drei unterschiedlichen Alpha-Synucleinopathien (PK, MSA und DLB) erweisen kann.

Beruhend auf den oben aufgeführten Ergebnissen vorangegangener Studien besteht die hier beantragte Untersuchung mittels einer Smartphone-App aus sieben kurzen, aktiven Tests zur 1) Stimmcharakteristik, 2) zum Gleichgewicht, 3) zum Gangverhalten, zum 4) Ruhe- und 5) Haltetremor der Hände sowie aus einem 6) Reaktions- und 7) Geschicklichkeitstest. Die Smartphone-App aktiviert, spezifisch für den jeweiligen Test, Sensoren (z.B. Bewegungssensor, Annäherungssensor oder Gyroskop), die handelsüblich bereits in jedem gängigen Smartphone enthalten sind. Keiner der Tests dauert länger als 45 Sekunden. Die Messung ist daher nicht invasiv, unkompliziert und dauert in der Regel nicht länger als ca. 7 Minuten. Außerdem erlaubt die Smartphone-App eine passive Messung von (alltäglichen) Bewegungsmustern durch oben aufgeführte Sensoren und dem GPS-System (Globales Positionsbestimmungssystem). Diese Messung läuft im Hintergrund und wird von den Patienten kaum bemerkt, d.h. die Patienten werden in ihren alltäglichen Handlungsabläufen nicht eingeschränkt.

Diese Smartphone-App wurde in den Vorläufer-Studien mit PK-Patienten bereits erfolgreich angewendet (14, 15, 16).

Es bleibt nun zu klären, inwiefern sich Veränderungen bei Sprechvorgängen, bei der Körperbewegung und der Geschicklichkeit mit der Smartphone-App bereits in der Prodromalphase RBD detektieren lassen und inwiefern die hier beantragten nicht invasiven,



kurzen, unkomplizierten und kostengünstigen Messungen mit den daraus erhobenen **objektiven** Parametern zur Früherkennung im echten Sinne, d.h. zur prodromalen Erkennung einer neurodegenerativen Erkrankungen bei RBD-Patienten, und zur therapiebegleitenden Verlaufskontrolle beitragen können.

Vor diesem Hintergrund ergeben sich für das hier beantragte Projekt der Querschnitt-Studie folgende *Forschungsfragen*:

Forschungsfrage 1: Unterscheiden sich RBD-Patienten in ihrer Stimmcharakteristik, in ihren Körperbewegungen und ihrer Geschicklichkeit von gesunden Kontrollpersonen, von Patienten mit einer Parkinson-Krankheit, einer Multi-System-Atrophie oder einer Progressiven Supranuklearen Blickparese ?

In Anlehnung an die hier beantragte Querschnitt-Studie lassen sich in einem zweiten Schritt im Rahmen der später geplanten Längsschnitt-Studie mit regelmäßigen Messwiederholungen potentiell folgende Forschungsfragen klären:

Forschungsfrage 2 (spätere Längsschnitt-Studie): Wie frühzeitig ist es möglich mit Hilfe von Smartphone-basierten Messungen motorische Auffälligkeiten und eventuelle, systematische Unterschiede bei Patienten mit einer RBD (Symptome typisch für Parkinson Krankheit versus Multi-System-Atrophie) zu detektieren?

Forschungsfrage 3 (spätere Längsschnitt-Studie): Lässt sich aus wiederholten Smartphone-erhobenen Parametern der Stimmcharakteristik, der Körperbewegung und Geschicklichkeit bzw. aus den daraus resultierenden, spezifischen behavioralen Mustern der weitere Krankheitsverlauf (schnelle oder langsame Progression in Richtung Konversion zu Parkinson-Krankheit oder Multi-System-Atrophie) bei Patienten mit einer RBD vorhersagen?

9. Art der Fragestellung

Es handelt sich um die Prüfung eines diagnostischen Verfahrens.



10. Studiendesign

Es handelt sich um eine experimentelle, nicht invasive Querschnitt-Studie (die dann gegebenenfalls in eine Langzeitstudie übergehen soll).

11. Zeitplan

Geplanter Studienbeginn ist September 2015; voraussichtliches Studienende der Querschnitts-Studie ist Februar 2017.

(Es ist dann längerfristig geplant, die Smartphone-basierten Messungen darauffolgend in 9-monatigen Abständen bei den Studienteilnehmern zu wiederholen. Die Langzeiterhebung wird mit diesem Ethikantrag nicht beantragt.)

12. Hypothesen

Hypothesen

Vor dem theoretischen Hintergrund ergeben sich für die oben aufgeführten Forschungsfragen folgende Hypothesen:

Hypothesen zur Forschungsfrage 1 (hier beantragte Querschnitt-Studie):

RBD-Patienten unterscheiden sich in den erhobenen Parametern zur Stimmcharakteristik, Körperbewegung und Geschicklichkeit von gesunden Kontrollpersonen.

Auffälligkeiten in den spezifischen Parametern sollten sich gemäß der langsamen Krankheitsprogression graduell zwischen den Gruppen aufzeigen. Das heißt, dass sich gemäß den oben erwähnten Vorstudien ein signifikanter Unterschied zwischen PK-Patienten und gesunden Kontrollpersonen zeigt, während sich bei RBD-Patienten als Prodromalstufe eher geringere Unterschiede in der Stimmcharakteristik, Veränderungen der Körperbewegung und in der Geschicklichkeit im Vergleich zu gesunden Kontrollgruppen aufzeigen (Gesunde < RBD < PK), die sich jedoch von gesunden Normwerten unterscheiden.

H1: PK-Patienten zeigen eine höhere Stimmfluktuation (Variabilität) als gesunde Kontrollpersonen.

H2: PK-Patienten zeigen höhere Stimmfluktuationen (Variabilität) als RBD-Patienten.



H3: RBD-Patienten zeigen höhere Stimmfluktuationen (Variabilität) als gesunde Kontrollpersonen, jedoch sind diese Stimmfluktuationen geringer ausgeprägt als bei PK-Patienten.

H4: PK-Patienten zeigen größere Bewegungsamplituden im Gleichgewichts-, Gang-, Ruhetremor und Haltetremortest als gesunde Kontrollpersonen.

H5: PK-Patienten zeigen größere Bewegungsamplituden im Gleichgewichts-, Gang-, Ruhetremor und Haltetremortest als RBD-Patienten.

H6: RBD-Patienten zeigen größere Bewegungsamplituden im Gleichgewichts-, Gang-, Ruhetremor und Haltetremortest) als gesunde Kontrollpersonen, jedoch sind diese Bewegungsamplituden geringer ausgeprägt als bei PK-Patienten.

H7: PK-Patienten zeigen längere Reaktionszeiten im Reaktions- und Geschicklichkeitstest als gesunde Kontrollpersonen.

H8: PK-Patienten zeigen längere Reaktionszeiten im Reaktions- und Geschicklichkeitstest als RBD-Patienten.

H9: RBD-Patienten zeigen längere Reaktionszeiten im Reaktions- und Geschicklichkeitstest als gesunde Kontrollpersonen, jedoch sind diese Reaktionszeiten kürzer ausgeprägt als bei PK-Patienten.

Hypothesen zur Forschungsfrage 2 (spätere Längsschnitt-Studie):

Der Sprechvorgang ist eine der kompliziertesten motorischen Leistungen des Menschen und es wird daher angenommen, dass gerade das Sprechen bereits bei einem relativ geringen Grad an nigrostriärer Degeneration geringe, aber bereits erfassbare Veränderungen aufweisen sollte. Zudem fanden Harel, Cannizzaro und Snyder (22) in ihren retrospektiven Analysen des „freien Sprechens“ mit Hilfe von archivierten Videoaufzeichnungen, dass sich eine reduzierte Frequenzvariabilität des Sprechens bereits bis zu 5 - 6 Jahre vor der Diagnosestellung einer Parkinson-Erkrankung zeigte. Dies und der zu erwartende Wissenszugewinn durch die hier beantragte Querschnitt-Studie bietet zu einem späteren Zeitpunkt die Grundlage zur konkreten Hypotheseneruierung.

Hypothesen zur Forschungsfrage 3 (spätere Längsschnitt-Studie):

Bei der dritten Forschungsfrage handelt es sich um eine rein explorative Fragestellung. Eine



derartige Untersuchung wurde unserem Wissen nach bis jetzt noch nicht durchgeführt.

13. Beschreibung des methodischen Vorgehens:

Messungen der Stimmcharakteristik, der Körperbewegung sowie der Geschicklichkeit mit Hilfe Smartphone-basierter, digitaler Aufzeichnungen

– a) im Klinik-Setting und b) im Alltags-Setting

a) Messbedingungen

Der Inhalt der von der Universität Oxford verwendeten Smartphone-App für Android-Handys wird für die Marburger Kohorte ins Deutsche übersetzt und ist so zusammengestellt, dass die zeitlichen Aufwendungen und allgemeinen Anforderungen an die Versuchspersonen / Patienten möglichst gering gehalten werden. Die Übersetzung der Instruktionen werden zuerst ins Deutsche übersetzt und dann im zweiten Schritt von einem ‚Native Speaker‘ wieder ins Englische rückübersetzt, damit sichergestellt werden kann, dass die Instruktionen identisch sind.

Je nach Wunsch der Studienteilnehmer kann die App entweder direkt auf das eigene Smartphone heruntergeladen und installiert werden oder es steht ihnen für den Messzeitpunkt ein von uns bereitgestelltes Smartphone mit vorinstallierter App zur Verfügung. Entscheidet sich der Studienteilnehmer dafür, sein eigenes Smartphone zu benutzen, wird ihm/ihr per E-Mail (Textentwurf liegt dem Ethikantrag im Anhang bei) der digitale Link zum Herunterladen der App (deutsche Version) zugesandt.

Die Smartphone-App beinhaltet

- I) sieben aktive, kurze Messungen der Stimmcharakteristik, Körperbewegung und Geschicklichkeit (im Weiteren „*aktive Mini-Tests*“ genannt) sowie
- II) eine passive, kontinuierliche Aufzeichnung (im Weiteren „*passives Monitoring*“ genannt) der Körperbewegung.

Im Rahmen der geplanten Studie werden die Versuchspersonen/Patienten gebeten, die „*aktiven Mini-Tests*“ zum einen

la) einmalig vor oder nach dem Kontrollbesuch bei Ihrem betreuenden Arzt in der Klinik für



Neurologie durchzuführen und zum anderen darauffolgend

IIb) eine Woche lang vier Mal täglich (in regelmäßigen Abständen über den Tag verteilt) selbstständige Messungen im Alltag zu tätigen.

Keiner der sieben „aktiven Mini-Tests“ dauert länger als 45 Sekunden. Ein kompletter Messdurchgang sollte – in Abhängigkeit von den selbstgewählten Pausen durch den Studienteilnehmer – nicht länger als circa fünf bis sieben Minuten dauern. Die reine Messzeit beträgt drei Minuten (fünfmal 20-Sekunden-Aufzeichnungen sowie zweimal 45-Sekunden-Aufzeichnungen). Neben den bereits in der Smartphone-App enthaltenen schriftlichen Instruktionen wird den Studienteilnehmern ein Informationsblatt mit Hinweisen und Erklärungen zur Durchführung der eigenständigen Testung zu Hause ausgehändigt (das Informationsblatt ist dem Ethikantrag im Anhang beigelegt).

Neben den „aktiven Mini-Tests“ werden die Studienteilnehmer in der gleichen Messwoche gebeten,

II) das „passive Monitoring“ so oft wie möglich während ihrer Wachstunden im Alltag einzuschalten und das Smartphone dafür am Körper zu tragen. Dazu verstauen sie dieses entweder in ihrer Hosen- oder Jackentasche oder benutzen einen von uns zur Verfügung gestellten Handy-Tragegurt. Auf diese Weise ist das „passive Monitoring“ kaum spürbar und es kann eine Beeinträchtigung der alltäglichen Tätigkeiten und Handlungsabläufe durch das Monitoring weitestgehend ausgeschlossen werden. Zudem werden die Studienteilnehmer darüber informiert, dass sie zu jeder Zeit selbst entscheiden, wann und wie lange sie bereit sind, die Aufzeichnung einzuschalten.

Im Folgenden Abschnitt b) sind die Messabläufe der einzelnen aktiven Tests und der passiven Aufzeichnung detailliert beschrieben:

b) Versuchsablauf

Zum ersten Messzeitpunkt erhalten die Studienteilnehmer zunächst eine ausführliche Aufklärung zum allgemeinen Ablauf der Studie sowie Erläuterungen zu den digitalen Aufzeichnungen im Rahmen der Smartphone-basierten Testungen und deren Durchführung.



Im Falle, dass die Nutzung des eigenen Handys gewählt wird, erhält der Studienteilnehmer eine E-Mail (vorgefertigte Fassung ist dem Ethikantrag im Anhang beigefügt), in der sich der Link und weitere Erklärungen zum Herunterladen und Installieren der Smartphone-App befinden. Sollte sich für ein von uns gestelltes Gerät entschieden werden, wird dem Probanden/Patienten die Handhabung des Leih-Smartphones ausführlich erklärt.

Bestehen im Weiteren keine zusätzlichen Fragen, wird nun gemeinsam mit den Studienteilnehmern die Smartphone-App geöffnet und im „Demo-Modus“ (dieser Modus der App dient zum Erklären und Ausprobieren der App; hier werden keine Daten aufgezeichnet) die Handhabung der einzelnen „aktiven Mini-Tests“ sowie das Ein- und Ausschalten des „passiven Monitorings“ besprochen.

1a/1b) Nach Öffnen der Smartphone-App werden die Studienteilnehmer durch Antippen des „aktive Tests“-Buttons zum ersten der insgesamt sieben Mini-Tests weitergeleitet.

Im Folgenden finden Sie die einzelnen „aktiven Mini-Tests“ aufgeführt und deren Ablauf im Detail beschrieben:

(1) Test zur Stimmcharakteristik

Bei dem ersten aktiven Mini-Test zur Stimmcharakteristik handelt es sich um eine ca. 20 sekundige Aufzeichnung der Stimme durch das im Smartphone befindliche Mikrofon. Bei diesem Test wird der Proband/Patient aufgefordert, die Messung durch Antippen eines sich auf dem Touchscreen befindlichen Button zu starten, das Handy dann – ähnlich wie beim Telefonieren – in die Nähe des Mundes zu führen, tief einzuatmen und dann – ähnlich wie beim Zahnarzt oder Hals-Nasen-Ohren-Arzt – so lange und so gleichmäßig wie möglich „aaaaah“ zu sagen. Die Aufzeichnung wird automatisch von der App beendet und mit einem Vibrationston kenntlich gemacht. Bei diesem Test wird dem Probanden/Patienten erläutert, dass es nicht notwendig ist, den Ton 20 Sekunden lang zu halten. Wünschenswert ist eine Aufzeichnung von mindestens sechs Sekunden.

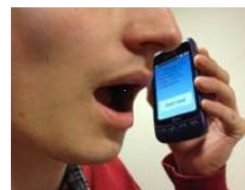


Abbildung 1



(2) Gleichgewichtstest

Bei dem zweiten aktiven Mini-Test zum Gleichgewicht wird der Proband/Patient gebeten, einen Ort aufzusuchen, an dem er freihändig und ohne Hilfe stehen kann. Nach Antippen des Start-Buttons auf dem Touchscreen wird der Bewegungssensor im Gerät aktiviert. Der Proband/Patient verstaut nun das Smartphone in seine/ihre Hosentasche (oder benutzt den Tragegurt). Die Aufzeichnung beginnt leicht zeitverzögert automatisch und wird durch einen Vibrationston kenntlich gemacht. Der Proband/Patient steht so lange aufrecht bis das Smartphone erneut vibriert und die Aufzeichnung nach 20 Sekunden beendet ist.

(3) Gangtest

Beim dritten aktiven Mini-Test zur Gangcharakteristik wird der Proband/Patient gebeten, einen Ort aufzusuchen, wo er ungehindert ungefähr 18 Meter ungestört geradeaus gehen kann. Nachdem er den Start-Button angetippt hat und der Bewegungssensor im Smartphone aktiviert ist, verstaut er/sie das Handy wieder in seine/ihre Hosentasche (oder benutzt den Tragegurt – siehe *Abbildung 2*). Die Aufzeichnung beginnt wiederum leicht zeitverzögert, wird mit einem Vibrationston angezeigt und der Proband/Patient geht nun in seinem eigenen Tempo geradeaus bis die Aufzeichnung durch einen erneuten Vibrationston nach 20 Sekunden automatisch beendet wird. Sollte die vorhandene Wegstrecke nicht ausreichend lang genug sein, kann der Proband während der Aufzeichnung umdrehen und die gewählte Strecke ein zweites Mal gehen.

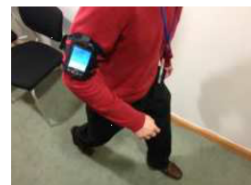


Abbildung 2

(4) Ruhetremortest

Beim vierten aktiven Test zum Ruhetremor wird der Proband/Patient aufgefordert, sich bequem aufrecht hinzusetzen, das Smartphone mit der vom Tremor am meisten betroffenen Hand festzuhalten (Patienten ohne merklichen Tremor benutzen zum Festhalten des Smartphones eine Hand ihrer Wahl) und dann beide Hände leicht auf den Oberschenkeln ruhen zu lassen (siehe *Abbildung 3*). Wenn der Proband/Patient soweit ist, wird durch



erneutes Antippen des Start-Buttons die Messung gestartet und dadurch der Bewegungssensor aktiviert. Danach soll der Proband/Patient die Augen schließen und von 100 rückwärts zählen (100 – 99 – 98 – und so weiter). Die Aufzeichnung startet wieder leicht zeitverzögert und wird mit einem Vibrationston kenntlich gemacht. Nach 45 Sekunden beendet die App mit einem erneuten Vibrationston die Aufzeichnung automatisch.

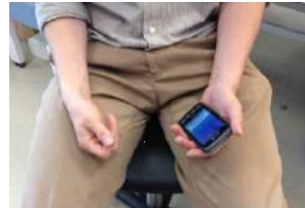


Abbildung 3

(5) Haltetremortest

Beim fünften aktiven Test zum Haltetremor wird der Proband/Patient erneut aufgefordert, sich aufrecht hinzusetzen und das Smartphone in die vom Tremor am meisten betroffene Hand festzuhalten (auch hier: Patienten ohne merklichen Tremor benutzen zum Festhalten des Smartphones eine Hand ihrer Wahl). Darüber hinaus wird er/sie gebeten, nach Antippen des Start-Buttons auf dem Display die Hand mit dem Smartphone nach vorne gerade auszustrecken. Die Aufzeichnung beginnt leicht zeitverzögert und wird durch einen ersten Vibrationston angezeigt. Der Proband/Patient hält den Arm so lange nach vorne gestreckt bis die App die Aufzeichnung nach 45 Sekunden durch einen erneuten Vibrationston beendet. Bei diesem Test wird dem Probanden/Patienten erklärt, dass er/sie versuchen soll, den Arm so lange wie möglich gestreckt zu halten, er/sie aber nicht die kompletten 45 Sekunden Aufzeichnungsdauer durchhalten muss.

(6) Reaktionstest

Beim sechsten aktiven Test zum Reaktionsvermögen wird der Proband/Patient gebeten, das Smartphone auf eine ebene Oberfläche zu legen. Wenn er/sie soweit ist, wird nach Antippen des Start-Buttons der Reaktionstest gestartet. In diesem soll der Studienteilnehmer mit dem Finger auf eine auf dem Bildschirm erscheinende viereckige Fläche tippen und diese so lange gedrückt halten, bis das Viereck wieder vom Bildschirm verschwindet. Erscheint das Viereck erneut, tippt der Proband/Patient wieder darauf und hält es mit dem Finger gedrückt und



zieht den Finger weg sobald das Viereck vom Bildschirm verschwindet. Der Berührungssensor des Smartphones erfasst in dieser 20-sekündigen Aufzeichnung die Reaktionszeiten der Studienteilnehmer von mehreren unterschiedlich langen Präsentationen des Vierecks.

(7) Geschicklichkeitstest

Beim letzten aktiven Test zur Geschicklichkeit wird der Proband/Patient wie schon im vorherigen Test gebeten, das Smartphone auf eine flache Oberfläche zu legen und durch Antippen des Start-Buttons die Aufzeichnung zu starten. Auf dem Bildschirm erscheinen nun sowohl unten links als auch rechts jeweils ein Rechteck. Der Studienteilnehmer wird gebeten abwechselnd mit dem Zeige- und Mittelfinger seiner/ihrer dominanten Hand möglichst in einem gleichmäßigen Rhythmus die beiden Rechtecke nacheinander anzutippen. Ähnlich wie beim Klavierspielen soll abwechselnd mit dem Zeigefinger das linke Rechteck auf dem Display und mit dem Mittelfinger das rechte Rechteck auf dem Display angetippt werden (siehe Abbildung 4). Die Aufzeichnung wird nach 20 Sekunden automatisch von der App beendet.



Abbildung 4

Mit Beenden des Geschicklichkeitstests ist die Messung abgeschlossen und die Smartphone-App kann geschlossen werden.

Der soeben beschriebene Messablauf mit den sieben „Mini-Tests“ wird einmalig in der Klinik im Beisein des Versuchsleiters durchgeführt **(Ia)**.

Im Anschluss daran werden die Studienteilnehmer gebeten eine Woche lang tägliche Messungen in ihrem Alltag eigenständig vorzunehmen **(Ib)**. Hierfür sollen sie in der einen Messwoche zuhause an sieben Tagen jeweils vier Messdurchläufe mit den 7 „Minitests“ (demnach 4 Messungen X 7 Tage = insgesamt 28 Messdurchläufe mit allen „Minitests“) in



regelmäßigen Abständen über den Tag verteilt durchführen.

Hinweis: In der späteren Langzeitstudie soll der oben ausführlich beschriebene Messblock (Ia - einmalig in der Klinik und Ib - eine Woche lang im Alltag zuhause) in 9-monatigen Abständen sechs Mal wiederholt werden. Diese Studie wird mit diesem Ethikantrag nicht beantragt.

II) Wie bereits erwähnt, werden die Studienteilnehmer zusätzlich gebeten, in der gleichen Messwoche im Alltag während ihrer Wachstunden das „*passive Monitoring*“ zu aktivieren. Dafür wird die Smartphone-App geöffnet und im Hauptmenü der „Monitoring starten“-Button angetippt. Mit Antippen des Buttons wird im Smartphone sowohl der im Gerät enthaltene Bewegungssensor, Annäherungssensor, Helligkeitssensor, Kompass als auch das GPS-System aktiviert. Mit Hilfe dieser Aufzeichnungen können differenzierte Bewegungsmuster (z.B. Gangbeschleunigungen bergauf oder bergab) erfasst und später pseudoanonymisiert ausgewertet werden (nähere Angaben zu den Parametern entnehmen Sie bitte dem nachfolgenden Abschnitt 13 c) „erhobene Parameter“). Das im Smartphone enthaltene Mikrofon wird durch die App zu keiner Zeit aktiviert.

Die Smartphone-App zeigt an, wann die Messung gestartet (Datum und Uhrzeit) wurde und wann sie wieder beendet wurde (Datum und Uhrzeit). Zudem informiert die App den Probanden/Patienten darüber, wann das Monitoring zuletzt benutzt wurde und zeigt auf, ob die Messung aktuell aktiviert („Messung ist aktiviert.“) oder deaktiviert („Messung ist deaktiviert.“) ist. Diese Informationen bieten dem Studienteilnehmer zu jeder Zeit die Möglichkeit zur Kontrolle über vergangene Messverläufe bzw. den aktuellen Messstatus.

Nach Ablauf der Messwoche gibt der Studienteilnehmer das Leihgerät entweder persönlich oder per vorfrankiertem Posteinwurf an uns zurück. Sollte das eigene Smartphone genutzt worden sein, steht es dem Studienteilnehmer frei, die App zu löschen.

c) Erhobene Parameter

Im *Stimmtest* werden unter anderem Parameter wie Tonlautstärke und Stabilität der Stimme bestimmt und ausgewertet.



Im *Gleichgewichtstest*, *Gangtest*, *Ruhetremor-* und *Haltetremortest* werden oszillatorische Parameter wie Schwankungsamplituden um die Gravitationsmitte (Bewegungsamplituden) erhoben und ausgewertet.

Im *Reaktions-* und *Geschicklichkeitstest* werden Reaktionszeiten bestimmt und ausgewertet. Im *Passiven Monitoring* werden Bewegungsamplituden, -geschwindigkeiten sowie die Häufigkeit von Bewegungen erfasst und analysiert.

14. Haupt- und Nebenzielkriterien

Hauptzielkriterien sind die Unterschiede der Messgrößen zur Stimmcharakteristik, Körperbewegung und Geschicklichkeit zwischen den einzelnen Untersuchungsgruppen. Nebenzielkriterien sind spezifische Muster, die sich aus der Kombination der erhobenen Hauptzielkriterien ergeben, d.h. es wird untersucht, inwiefern sich pro Studiengruppe charakteristische Muster aus den erhobenen Messgrößen ergeben und inwiefern sich diese zwischen den Untersuchungsgruppen unterscheiden.

15. Patienten- / Probandenzahl

Am Standort Marburg ist es geplant pro Untersuchungsgruppe N = 35 Patienten mit RBD, PK, MSA und PSP einzuschließen. Außerdem sollen 4 x N = 35 gesunde Kontrollpersonen rekrutiert werden, die in der Gruppe bezüglich des Alters und des Geschlechts mit den Patientengruppen gematcht werden sollen („matched controls“).

Fallzahlberechnung und Fallzahlbegründung

Hier beantragte Querschnitt-Studie: Die Berechnung des Umfangs der Stichprobe basiert auf den Annahmen der erwarteten Effektstärke (die sich aus vorangegangenen Studien mit PK-Patienten ergibt) bei einem angelegten Signifikanzniveau von 5% und einer Testpower von 80%. Aus diesen Annahmen ergibt sich anteilig für die Kohorte am Standort Marburg eine Stichprobengröße von N = 35 pro Patientengruppe.

Mögliches Folgeprojekt - Langzeit-Studie: Wir erwarten aufgrund der Erfahrungen aus vorangegangenen experimentellen Langzeitstudien eine Dropout-Rate von mind. 30 % und einen technischen Datenverlust von ca. 15%.



Die Marburger Querschnitt-Studie ist Teil einer bizenrischen Studie – in Kooperation mit der klinischen Arbeitsgruppe von Dr. Hu, University Oxford, United Kingdom.

Nach den ersten 4 Monaten der Querschnitts-Studie in Marburg wird eine Analyse der Rekrutierungssituation durchgeführt. Danach wird der Antrag auf eine – auf der Querschnitts-Studie basierenden – Längsschnitt-Studie gestellt:

Um mögliche Drop-outs zum 6. Messzeitpunkt der Langzeit-Studie kompensieren zu können, ist geplant, insgesamt $n = 120$ Personen pro Patientengruppe zu erheben. Diese Stichprobengröße bezieht sich auf beide Erhebungskohorten zusammen, also Erhebungen hier in Marburg und Erhebungen in Oxford, Vereinigtes Königreich.

Wie bereits erwähnt ist es geplant am Standort Marburg anteilig insgesamt $N = 35$ pro Untersuchungsgruppe zu erheben.

16. Besonders schutzbedürftige Personen

Die Studie wird ausschließlich an einwilligungsfähigen Erwachsenen durchgeführt. Kinder und nicht einwilligungsfähige Erwachsene sind ausgeschlossen.

17. Ein- / und Ausschlusskriterien Patienten bzw. Probanden

Einschlusskriterien

Kontrollgruppe

- Männliche und weibliche Personen im Alter zwischen 18 und 80 Jahren.
- Schriftliches Einverständnis zur freiwilligen Teilnahme an der Studie.
- Fähigkeit, Risiken und Komplikationen der Untersuchungen zu verstehen und zu beurteilen.

Patienten mit RBD

- Männliche und weibliche Personen im Alter zwischen 18 und 80 Jahren.
- Diagnose RBD entsprechend den Kriterien der International Classification of Sleep Disorders [ASDA Criteria 2005].
- Schriftliches Einverständnis zur freiwilligen Teilnahme an der Studie.
- Fähigkeit, Risiken und Komplikationen der Untersuchungen zu verstehen und zu beurteilen.



beurteilen.

Patienten mit PK

- Männliche und weibliche Personen im Alter zwischen 18 und 80 Jahren.
- Diagnose einer PK gemäß Leitlinien der Deutschen Gesellschaft für Neurologie.
- Schriftliches Einverständnis zur freiwilligen Teilnahme an der Studie.
- Fähigkeit, Risiken und Komplikationen der Untersuchungen zu verstehen und zu beurteilen.

Patienten mit MSA

- Männliche und weibliche Personen im Alter zwischen 18 und 80 Jahren.
- Diagnose einer wahrscheinlichen MSA gemäß Leitlinien der Deutschen Gesellschaft für Neurologie.
- Schriftliches Einverständnis zur freiwilligen Teilnahme an der Studie.
- Fähigkeit, Risiken und Komplikationen der Untersuchungen zu verstehen und zu beurteilen.

Patienten mit PSP

- Männliche und weibliche Personen im Alter zwischen 18 und 80 Jahren.
- Diagnose einer wahrscheinlichen PSP gemäß Leitlinien der Deutschen Gesellschaft für Neurologie.
- Schriftliches Einverständnis zur freiwilligen Teilnahme an der Studie.
- Fähigkeit, Risiken und Komplikationen der Untersuchungen zu verstehen und zu beurteilen.

Ausschlusskriterien

- Nicht einwilligungsfähige Personen.
- Patienten mit einem Medikamentös-induzierten Parkinson-Syndrom oder einem Parkinson-Syndrom, das eine bekannte Ursache aufweist (z.B. subcortikale arteriosklerotische Enzephalopathie)
- Kinder sowie Personen unter 18 Jahren und über 80 Jahren.



18. Darlegung des statistischen Auswertverfahrens

Aus den vorausgegangenen Untersuchungen mit PK-Patienten und gesunden Kontrollgruppen ist davon auszugehen, dass die ausgewählten Messzeiten (zwischen 20 und 45 Sekunden pro „Mini-Test“) eine hinreichende Anzahl an Datenpunkten ergeben, um auf individueller Ebene akustische, oszillatorische sowie zeitliche Daten für die Analyse zu erhalten.

Die Daten der spezifischen Kenngrößen zur Stimmcharakteristik, Körperbewegung und Geschicklichkeit werden mittels MATLAB (The MathWorks INC., USA) unter Verwendung spezifisch entwickelter, mathematischer Algorithmen von der AG um Dr. Max Little ermittelt. Die statistische Auswertung erfolgt ebenfalls mit MATLAB.

19. Darlegung und Angabe aller projektbedingten Handlungen am Patienten oder Probanden

Kontakt zum Patienten/Probanden besteht während der Aufklärung über die Studie durch die Versuchsleitung und zur Durchführung der ca. 7-minütigen Messung der „aktiven Tests“ mit Hilfe der Smartphone-App während eines Kontrollbesuches hier in der Klinik. Die regelmäßigen Messungen im Alltag innerhalb der darauffolgenden Messwoche führt der Proband / Patient eigenständig zu Hause durch.

20. Abbruchkriterien

Ein Abbruch ist bei Willen des Patienten bzw. Probanden jederzeit möglich, wobei für diesen hierdurch keinerlei Nachteile entstehen. Weiterhin kann die Studie aus gesundheitlichen, technischen oder anderen Gründen durch die Versuchsleiter abgebrochen werden.

21. Art der Dokumentation

Mit der Smartphone-App werden keine personenbezogenen Daten erhoben. Die erhobenen Zielparameter werden elektronisch erfasst und verschlüsselt auf einem projektbezogenen, gesicherten Server übermittelt und gespeichert. Informationen zur Entschlüsselung der pseudonymisierten Daten stehen nur projektbezogenen wissenschaftlichen Mitarbeitern zur Verfügung.



22. Angabe der Anzahl konkurrierender Studien / Projekte

Wir führen derzeit keine Studien durch, die die Motorik der Patienten mit RBD, PK, MSA oder PSP oder gesunder Kontrollen erfasst.

Parallel werden nicht-invasive Studien zur Pupillendilatation und Ausatemluft und eine minimal invasive Studie mit Hautbiopsie diesen Patientengruppen angeboten.

23. Rekrutierung

Die meisten Personen, die an einer RBD leiden, melden sich aus dem ganzen Bundesgebiet per Brief, email oder Telefon bei unseren Studienkoordinatoren (Frau Baumgardt und Frau Siepe) von selbst, das heißt freiwillig. Die Rekrutierung der Kontrollgruppe erfolgt durch mündliche Anfrage bei Patienten des interdisziplinären Schlaflabors (Leitung Prof. Dr. Koehler, Klinik für Pulmologie, Marburg) und Angehörigen von Patienten in den Spezialprechstunden der Klinik für Neurologie. Die Rekrutierung der Patienten mit PK, MSA oder PSP erfolgt aus den Spezialambulanzen der Klinik für Neurologie in Marburg nach Kontaktaufnahme durch die betreuenden Ärzte und Ärztinnen.

24. Aufwandsentschädigung für Patienten / Probanden

Eine finanzielle Aufwandsentschädigung ist nicht vorgesehen.

Reisekosten und ggf. Kosten für Übernachtung (siehe Absatz 23 – Einzugsgebiet für betroffene Personen) werden erstattet.

C Aufklärung und Zustimmung

25. Aufklärungsblatt, Einverständniserklärung

Patienten, welche die Einschlusskriterien erfüllen, werden durch die Versuchsleitung über die Studie informiert und bekommen ein Patienten- bzw. Probandeninformationsblatt ausgehändigt. Nach der mündlichen und schriftlichen Aufklärung können die Probanden / Patienten ihre schriftliche Einwilligung zur Teilnahme an der Studie geben.

Das Informationsblatt und die Einverständniserklärung sind dem Ethikantrag im Anhang beigelegt.



26. Wer informiert / klärt auf?

Die Patienten- bzw. Probandeninformation und die Ausgabe eines Informationsblattes zur selbstständigen Durchführung der Tests im Alltag (die Formulare sind dem Ethikantrag im Anhang beiliegend) erfolgt durch die Versuchsleitung (Prof. Dr. Dr. W. Oertel, PD Dr. C. Best, Dr. D. Vadasz und Dipl.-Psych. K. Wilhelm). Zusätzlich besteht die Möglichkeit, bei Rück- und Nachfragen eine auf dem Informationsblatt aufgeführte Telefonnummer der Klinik für Neurologie anzurufen.

D Datenschutz

27. Datensammlung – Pseudonymisierung

Alle von der Smartphone-App erhobenen Daten sind nur anhand der Teilnehmer-Nummer identifizierbar, so dass sichergestellt wird, dass die Identität des Patienten bzw. des Probanden anonym bleibt. Die Kodierung der Teilnehmer-Nummer erfolgt anhand einer dreistelligen Zahlenfolge, die mit jedem zusätzlichen Patienten/Probanden um eins erhöht wird. Die Teilnehmer-Nummer wird außerdem mit einem Zusatz des Prefix „MR“ versehen, um die Teilnehmer-Nummern der Marburger Stichprobe zuzuordnen.

Die Verbindung zwischen den erhobenen Messdaten und der Teilnehmeridentität kann nur mit einer Schlüsseliste hergestellt werden. Diese wird getrennt von den Messdaten aufbewahrt und zwei Jahre nach Abschluss der Untersuchung (Querschnitts-Studie) vernichtet.

Die Teilnehmer werden schriftlich darüber aufgeklärt, dass ihre Daten pseudonymisiert und analysiert werden und dass eine vertrauliche Behandlung aller persönlichen Daten entsprechend der aktuell gültigen Datenschutzbestimmungen erfolgt.

Bei dem möglichen Folgeprojekt der Langzeitstudie wird zum gegebenen Zeitpunkt analog zum oben beschriebenen Vorgehen eine erneute, schriftliche Aufklärung zur Datensammlung/Pseudonymisierung durch die Versuchsleitung durchgeführt.

28. Speicherung und Übermittlung

Die pseudonymisierten Daten werden vom Smartphone aufgezeichnet und verschlüsselt via Internetverbindung auf den gesicherten, projektbezogenen Server übermittelt. Sobald sich



das Smartphone in der Nähe eines verfügbaren WLAN-Zugangs befindet, wird die App alle 15 Minuten automatisch einen Upload durchführen und die bis dahin erhobenen Daten auf den gesicherten Server übermitteln. Die Übermittlung der Daten kann bei Bedarf auch manuell durch den Studienteilnehmer durchgeführt werden, indem dieser im Haupt-Menü der Smartphone-App auf den Menüpunkt „Hochladen“ tippt. Die wissenschaftlichen Mitarbeiter in Oxford stellen mit kryptologischen Maßnahmen sicher, dass nur die mit dem Projekt betreuten Personen Zugriff auf die erhobenen Daten erhalten. Die Verbindung zwischen Messdaten und der Teilnehmeridentität kann nur mit einer Schlüsselliste hergestellt werden. Die Schlüsselliste der Marburger Kohorte verbleibt zu jeder Zeit in der Klinik für Neurologie (Verwahrung bei Elisabeth Sittig-Wiegand, Studienschwester) und kann von Mitgliedern anderer Institute nicht eingesehen werden.

29. Information des Hausarztes / Schweigepflichtentbindung

Dritte Personen werden nicht informiert.



Literaturverzeichnis

- (1) Ohayon, M.M. & Schenck, C.H. (2010). Violent behavior during sleep: prevalence, comorbidity and consequences. *Sleep Medicine*, 11: 941 – 946.
- (2) Iranzo, A., Tolosa E., Gelpi, E., Molinuevo, J.L., Valdeoriola, F., Serradell, M., Sanchez-Valle, R., Vilaseca, I., Lomena, F., Vilas, D., Llado, A. & Santamaria J. (2013). Neurodegenerative disease status and post-mortem pathology in idiopathic rapid-eye-movement sleep behaviour disorder: an observational cohort study. *Lancet Neurology*, 12: 443-53.
- (3) Iranzo, A., Santamaria, J. & Tolosa, E. (2009b). The clinical and pathophysiological relevance of REM sleep behavior disorder in neurodegenerative diseases. *Sleep Medicine Review*, 13: 385-401.
- (4) Schenck, C.H., Boeve, B.F. & Mahowald, M.W. (2013). Delayed emergence of a parkinsonian disorder or dementia in 81% of older men initially diagnosed with idiopathic rapid eye movement sleep behavior disorder: a 16-year update on a previously reported series. *Sleep Medicine*. Aug; 14,8: 744-748. doi: 10.1016/j.sleep.2012.10.009. Epub 2013, Jan 22.
- (5) Postuma, R.B., Aarsland, D., Barone, P., Burn, D.J., Hawkes, C.H., Oertel, W.H. & Ziemssen, T. (2012). Identifying Prodromal Parkinson's disease: Pre-Motor Disorders in Parkinson's disease. *Movement Disorders*, 27, 5: 617-626.
- (6) Dorsey, E.R., Constantinescu, R., Thompson, J.P. , Biglan, K.M., Holloway, R.G., Kieburtz, K., Marshall, F.J., Ravina, B.M., Schifitto, G. Siderowf, A. and Tanner, C.M. (2007). Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. *Neurology*, 68: 384–386.
- (7) Von Campenhausen, A.S., Bornschein, B., Wick, R., Bötzel, K., Sampaio, C., Poewe, W, Oertel, W., Siebert, U., Berger, K. & Dodel, R. (2005). Prevalence and incidence of Parkinson's disease in Europe. *European Neuropsychopharmacology*, 15: 473 – 490.
- (8) Fearnley, J.M. & Lees, A.J. (1991). Ageing Parkinson's disease: Substantia nigra regional selectivity. *Brain*, 114(5): 2283-2301.
- (9) Kordower, J.H., Olanow, C.W., Dodiya, H.B., Chu, Y., Beach, T.G., Adler, C.H., Halliday, G.M. & Bartus, R.T. (2013). Disease duration and the integrity of the nigrostriatal system in Parkinson's disease. *Brain*, 136:2419-2431.



- (10) Santamaria, J., Tolosa, E. & Valles, A. (1986). Parkinson's disease with depression: a possible subgroup of idiopathic Parkinsonism. *Neurology*, 36: 1130-1133.
- (11) Schuurman, A.G., van den Akker, M., Ensink, K.T., Metsemakers, J.F.M., Knottnerus, J.A., Leentjens, A.F.G. & Buntinx, F. (2002). Increased risk of Parkinson's disease after depression: A retrospective cohort study. *Neurology*, 58: 1501-1504.
- (12) Abbott, R.D., Petrovitch, H., White, L.R., Masaki, K.H., Tanner, C.M., Curb, J.D., Grandinetti, A., Blanchette, P.L., Popper, J.S. & Ross, G.W. (2001). Frequency of bowel movements and the future risk of Parkinson's disease. *Neurology*, 57: 456-462.
- (13) Katzenschlager, R. & Lees, A.J. (2004). Olfaction and Parkinson's syndromes: its role in differential diagnosis. *Current Opinion in Neurology*, 17: 417-423.
- (14) Arora, S., Little, M.A., Venkataraman, V., Donohue, S., Biglan, K. & Dorsey, E.R. (2013). High-accuracy Discrimination of Parkinson's Disease Participants from Healthy Controls Using Smartphones. *Movement Disorders*, 28(10): e11.
- (15) Tsanas, A., Little, M.A., McSharry, P.E., Spielman, J. & Ramig, L.O. (2012). Novel speech signal processing algorithms for high-accuracy classification of Parkinson's Disease, *IEEE Transactions on Biomedical Engineering*, 59(5): 1264-1271.
- (16) Tsanas, A., Little, M.A., McSharry, P.E. & Ramig, L.O. (2009). Accurate telemonitoring of Parkinson's disease progression by non-invasive speech tests. *IEEE Transactions on Biomedical Engineering*, 57(4): 884-893.
- (17) Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease (2003). State of the Art Review: The Unified Parkinson's Disease Rating Scale (UPDRS): Status and Recommendations. *Movement Disorders*, 18(7): 738-750.
- (18) Brunato, M., Battiti, R., Pruitt, D. & Sartori, E. (2013). Supervised and unsupervised machine learning on detection, monitoring and management of Parkinson's disease from passive mobile phone data, Michael J. Fox Foundation Kaggle Competition entry, LIONSolver Inc.
- (19) Liddle, J., Ireland, D., McBride, S.J., Brauer, S.G., Hall, L.M., Ding, H., Karunanithi, M., Hodges, P.W., Theodoros, D., Silburn, P.A. & Chenery, H.J. (2014). Measuring the Lifespace of People with Parkinson's Disease using Smartphones: Proof of Principle. *Journal of Medical Internet Research mHealth uHealth*, 2(1): e13.




- (20) Rusz, J., Bonnet, C., Klempir, J., Tykalova, T., Baborova, E., Novotny, M., Rulseh, A. & Ruzicka, E. (2015). Speech disorders reflect differing pathophysiology in Parkinson's disease, progressive supranuclear palsy and multiple system atrophy. *Journal of Neurology*: DOI 10.1007/s00415-015-7671-1.
- (21) Tsanas, A., Little, M.A., McSharry, P.E. & Ramig, L.O. (2010). Nonlinear analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity. *Journal of the Royal Society Interface*, 8(59): 842-855.
- (22) Harel, B., Cannizzaro, M. & Snyder, P.J. (2004). Variability in fundamental frequency during speech in prodromal and incipient Parkinson's disease: A longitudinal case study. *Brain and Cognition*, 56: 24-29.
- (23) Postuma, R.B., Lang, A.E., Gagnon, J.F., Pelletier, A. & Montplaisir, J.Y. (2012). How does parkinsonism start? Prodromal parkinsonism motor changes in idiopathic REM sleep behavior disorder. *Brain*, 135, 1860-1870.

Unterschrift

Prof. Dr. med. Dr. h.c. Wolfgang H. Oertel

PD Dr. med. Christoph Best

10.4 Votum der Ethikkommission



Philipps Universität
Marburg Marburg

Philipps-Universität - 35032 Marburg

Herrn
Prof. Dr. W. H. Oertel
Philipps-Universität Marburg
Fachbereich Medizin
Klinik für Neurologie
Baldingerstraße
D-35033 Marburg

Fachbereich Medizin
Dekanat/Ethikkommission
Prof. Dr. med. Gerd Richter (Vors.)

Tel.: 06421 586 6487
Fax: 06421 586 6585
Sek.: D. Raiss
E-Mail: ethikkom@staff.uni-marburg.de
Anschrift: Baldingerstrasse/Postfach 2360
35032 Marburg
Web: www.med.uni-marburg.de/ethikkomm
Az.: 116/15

Marburg, den 27.10.2015

Studie: „Smartphone-basierte Untersuchung zur Stimmcharakteristik, Körperbewegung und Geschicklichkeit bei Patienten mit REM-Schlaf-Verhaltensstörung (RBD), Parkinson-Krankheit (PK) und Multi-System-Atrophie (MSA), bei einer Krankheitskontrolle (progressive supranukleäre Blickparese (PSP)) und bei gesunden Kontrollpersonen.“

Eingereichte Unterlagen:

1. Ihre Stellungnahme vom 14.09.2015 (hier eingegangen am 23.10.2015) zu unserem Gutachten vom 27.08.2015,
2. Ethikantrag Version 14.09.2015,
3. Health Research Authority NRES Committee South Central – Berkshire v. 22.1..2013,
4. Information für Studienteilnehmer Version 14.09.2015,
5. Einverständniserklärung Version 14.09.2015,
6. Informationsblatt zur Durchführung der Smartphone-basierten Messungen zu Hause, Version 14.09.2015,
7. e-mail_zum_Herunterladen_der_App_Version_1Stand_09.06.2015.

Sehr geehrter Herr Prof. Oertel,

vielen Dank für Ihre ausführliche Stellungnahme vom 14.09.2015 zu unserem Gutachten vom 27.08.2015. Mit Ihrer Stellungnahme und den daraus resultierenden Korrekturen bzw. Modifikationen sind alle Auflagen erfüllt. Es bestehen nunmehr keine berufsethischen oder berufsrechtlichen Bedenken gegen die Durchführung der Studie.

Damit schließt die Ethikkommission des Fachbereichs Medizin der Philipps-Universität Marburg ihr Gutachten mit einem

positiven Ethikkommissionsvotum

ab.

Entsprechend der ausschließlich beratenden Funktion der Ethik-Kommission betrifft dieses Votum nur die ethische Beurteilung der Konzeption, der vorgesehenen Methoden, der Durchführung und Überwachung des betreffenden Projektes sowie der beabsichtigten Patientenaufklärung. Die ärztliche und juristische Verantwortung verbleibt jedoch uneingeschränkt beim Projektleiter und seinen Mitarbeitern.

Sekretariat : Frau Raiss Montag – Donnerstag 8.00 – 12.00 Uhr, Freitag 8.00 – 11.00 Uhr
Frau Becken Montag – Donnerstag 14.00 – 16.00 Uhr

- 2 -

Bitte geben Sie uns jede Änderung in der Protokolldurchführung bekannt. Es muss dann geklärt werden, ob das Votum der Ethik-Kommission noch Gültigkeit hat.

Über alle schwerwiegenden oder unerwarteten unerwünschten Ereignisse, die während der Studie auftreten und die Sicherheit der Studienteilnehmer oder die Durchführung der Studie beeinträchtigen könnten, muss der Vorsitzende der Ethik-Kommission unterrichtet werden.

Die Ethik-Kommission des Fachbereichs Humanmedizin der Philipps-Universität Marburg arbeitet gemäß den nationalen gesetzlichen Bestimmungen und den ICH-GCP-Richtlinien.

Außerdem benötigt die Ethik-Kommission einen Bericht nach Abschluss der Studie.

Mit freundlichen kollegialen Grüßen
für die Ethik-Kommission des
Fachbereichs Humanmedizin
der Philipps-Universität Marburg

Prof. Dr. med. G. Richter
(Vorsitzender Ethikkommission)

Verzeichnis der akademischen Lehrer und Lehrerinnen

Meine akademischen Lehrerinnen und Lehrer an der Philipps-Universität Marburg waren in alphabetischer Reihenfolge die Damen und Herren:

Adamkiewicz, Bartsch, Baum, Becker, Best, Czubayko, Daut, Decher, Dodel, Feuser, Fritz, Fuchs-Winkelmann, Grundmann, Gress, Hertl, Hofmann, Hoyer, Hundt, Kinscherf, Kircher, Lang, Leonhardt, Lill, Lohoff, Maier, Moll, Moosdorf, Mueller, Neubauer, Neumüller, Nimsky, Oberwinkler, Oertel, Oliver, Opitz, Pagenstecher, Plant, Preisig-Müller, Renz, Richter, Ruchholtz, Schäfer, Schieffer, Seitz, Thieme, Timmermann, Vogelmeier, Wagner, Weber, Weihe, Werner, Westermann, Wilhelm, Wrocklage, Wulf

Danksagung

Großer Dank gebührt Prof. Dr. Dr. h.c. Wolfgang Oertel für die Überlassung des Themas, die Bereitstellung der Mittel und der Rahmenbedingungen für eine interessante Dissertation mit vielen Einblicken in die Forschungsarbeit der gesamten Arbeitsgruppe, aber auch der klinischen Arbeit in der Neurologie. Weiterhin danke ich herzlich für die, insbesondere in der Endphase, enge und konstruktive Betreuung.

Frau Dr. Annette Janzen möchte ich auch ganz herzlich danken für die viele investierte Zeit, die immer rasche Hilfe und sehr konstruktive Zusammenarbeit, sowie die moralische Unterstützung. Die ausdauernde Unterstützung und Kontinuität haben die Arbeit maßgeblich mitgetragen.

Den Mitgliedern der Arbeitsgruppe danke ich für ihre Hilfsbereitschaft, die gute Stimmung, Freundlichkeit und fortwährende Unterstützung. Insbesondere die Unterstützung von Frau Elisabeth Sittig war von unschätzbarem Wert.

Mein ausdrücklicher Dank gilt ebenfalls meinen Freunden für die moralische Unterstützung und das gegenseitige Antreiben. Unser aller Ausdauer hat hiervon sicherlich sehr profitiert. Insbesondere hervorzuheben ist hier Patrick Schulze für seinen tatkräftigen Beistand.

Ganz besonders möchte ich mich bei meinen Eltern Corinna Beilharz und Werner Thal bedanken. Ohne ihre Unterstützung wäre die Arbeit nicht denkbar gewesen. Ich bedanke mich von ganzem Herzen für die Beratung, die Motivation, das Rücken-Freihalten, die Kinderbetreuung, um mir das Schreiben zu ermöglichen und vieles mehr.

Zu guter Letzt möchte ich mich bei meiner Frau Fritzi bedanken. Ihre unermüdliche und ausdauernde Unterstützung, das Ertragen meiner Klagen und ihre aufbauenden Worte haben mich durchhalten lassen. Sie hat mir den Rücken freigehalten, mich motiviert und angetrieben. Für ihre Hilfe bin ich unendlich dankbar.