

# Kurzbericht ParKInsonPredict

**Zuwendungsempfänger:** Technische Universität Dresden - Medizinische Fakultät Carl Gustav Carus - Klinik und Poliklinik für Neurologie, Förderkennzeichen: 16DKWN113A

**Vorhabenbezeichnung:** ParKInsonPredict – Entwicklung und Testen eines KI-Modells zur Vorhersage von Krankheitsprogressions-Mustern bei neurodegenerativen Erkrankungen am Beispiel der Parkinsonerkrankung auf Basis von multi-modalen Daten verschiedener Patientenkohorten.

**Laufzeit des Vorhabens / Berichtszeitraum:** 1. September 2022 – 30. Juni 2024

## Aufgabenstellung

Die übergeordnete Fragestellung von ParKInsonPredict war es, zu untersuchen, inwieweit KI-gestützte Algorithmen eine zuverlässige Vorhersage von Krankheitsprogressionsmustern bei der Parkinson-Krankheit ermöglichen. Dazu sollte in einem ersten Schritt aufbauend auf Vorarbeiten des Fraunhofer-Instituts für Algorithmen und Wissenschaftliches Rechnen (SCAI) ein KI-Modell zur Identifikation und Vorhersage von Parkinson-Verlaufstypen anhand multimodaler Daten der PPMI-Studie (Parkinson's Progression Markers Initiative) entwickelt und evaluiert werden. In einem zweiten Schritt sollte das bereits vortrainierte Modell auf verschiedene verfügbare Studiendatensätze angewendet werden. Auf diese Weise sollte evaluiert werden, inwieweit ein solches KI-Modell auf andere klinische Szenarien generalisierbar ist und somit die Generalisierbarkeit der gewonnenen Erkenntnisse demonstriert werden, was eine wesentliche Voraussetzung für den zukünftigen Einsatz in der Patientenversorgung darstellt. Ein weiterer Schwerpunkt war die detaillierte klinische Charakterisierung der so identifizierten Progressionsmuster und deren Vorhersage, wobei auch hier die Generalisierbarkeit dieser Erkenntnisse über verschiedene Studien hinweg evaluiert werden sollte.

## Wissenschaftlicher und technischer Stand, an den angeknüpft wurde

Künstliche Intelligenz hat in den letzten Jahren in der Medizin und insbesondere im Bereich neurodegenerativer Erkrankungen wie der Parkinson-Krankheit massiv an Bedeutung gewonnen und steht zunehmend im Fokus vieler aktueller Forschungsprojekte. Bisherige Forschungsarbeiten konzentrierten sich fast ausschließlich auf die Erfassung und Graduierung motorischer Symptome mittels verschiedener Sensoren und auf spezifische diagnostische Fragestellungen wie (I) die Differenzierung zwischen Patienten mit Parkinson-Erkrankung und gesunden Kontrollprobanden sowie (II) die Differenzierung zwischen Patienten mit Parkinson-Erkrankung und Patienten mit anderen Erkrankungen wie z.B. atypischen Parkinson-Syndromen.

Aus klinischer Sicht spielt nach der Diagnosestellung insbesondere die Früherkennung neu auftretender motorischer und nicht-motorischer Parkinson-Symptome eine entscheidende Rolle, um frühzeitig eine Therapie einleiten zu können. Im Bereich der Vorhersage des Krankheitsverlaufs gibt es bisher nur wenige Pilotstudien. Nur ein Teil der Studien nutzt direkt longitudinale Daten, um klinische Verlaufstypen zu identifizieren. Zudem basieren diese Studien ausschließlich auf Daten der PPMI-Studie, d.h. Patienten mit neu diagnostizierter Parkinson-Erkrankung (Diagnose < 2 Jahre, keine medikamentöse Therapie). Die Anwendbarkeit der verwendeten KI-Modelle auf Patienten mit einer bereits länger als 2 Jahre bestehenden Parkinson-Erkrankung, was klinisch die deutlich häufigere Situation darstellt, bleibt unklar.

## Ablauf des Vorhabens

Grundlage des beantragten Projektes war der am Fraunhofer-Institut für Algorithmen und Wissenschaftliches Rechnen (SCAI) entwickelte KI-Algorithmus VaDER (Variational Deep Embedding with Recurrence), der bereits in der Vergangenheit auf dem PPMI-Datensatz zur Identifikation und Vorhersage von klinischen Progressionsmustern trainiert

wurde. Im Arbeitspaket 1 sollte dieses Modell auf mehrere weitere Datensätze angewendet werden, so dass für jeden Patienten eine Zuordnung zu einem klinischen Verlaufstyp erfolgen kann. In Arbeitspaket 2 sollten die einzelnen Charakteristika (Baseline- und Verlaufskarakteristika) der identifizierten Verlaufstypen mittels statistischer Analysen untersucht werden. Durch die Anwendung auf Datensätze mehrerer klinischer Studien wurde eine höhere Robustheit der identifizierten Charakteristika der Verlaufstypen im Vergleich zu Vorarbeiten erwartet. Die Generalisierbarkeit der identifizierten Verlaufstypen sollte dabei durch einen statistischen Vergleich der identifizierten Charakteristika der Verlaufstypen zwischen den verschiedenen Studien erfolgen. In Arbeitspaket 3 sollte aufbauend auf den Vorarbeiten des SCAI ein Klassifikator zur Vorhersage der Verlaufstypen für die genannten Datensätze entwickelt und dabei die Generalisierbarkeit eines solchen Klassifikators evaluiert werden. Arbeitspaket 4 schließlich widmete sich der Veröffentlichung und Verbreitung der Ergebnisse sowie der Planung von Folgeprojekten.

## **Wesentlichen Ergebnisse sowie Zusammenarbeit mit anderen Forschungseinrichtungen**

In jeder der untersuchten Studien konnte ein schnell und ein langsam progredienter Verlaufstyp der Parkinson-Krankheit identifiziert werden. Die Verlaufstypen unterschieden sich hinsichtlich der Progression verschiedener klinischer Charakteristika: Motorische Symptome, nicht-motorische Symptome, bildgebende Marker (DaTSCAN), digitale Ganganalyse, Mortalität, und Therapieansprechen. Auch demographische und „externe Faktoren“ unterschieden sich statistisch und klinisch signifikant zwischen den beiden Verlaufstypen. Der schnell progrediente Verlaufstyp zeigte eine schnellere Progression der motorischen Symptome, insbesondere der Gangstörung, der nicht-motorischen Symptome (z.B. autonome Funktionsstörungen, Demenz) und eine erhöhte Mortalität. Mittels bildgebender Verfahren (DaTSCAN) konnte gezeigt werden, dass der schnell progrediente Verlaufstyp auch objektiv eine schnellere Neurodegeneration der Substantia nigra aufweist. Darüber hinaus war dieser Verlaufstyp durch ein schlechteres Ansprechen auf die medikamentöse Therapie gekennzeichnet. „Externe Faktoren“ wie Bildung, Körpergröße, Körpergewicht und Alter bei Diagnosestellung hatten ebenfalls einen signifikanten Einfluss auf den Krankheitsverlauf. Andere Faktoren wie Komorbiditäten und Begleitmedikation zeigten keinen Einfluss auf den Verlaufstyp.

Entsprechende Klassifikatoren zur Vorhersage des Verlaufstyps für einzelne Patienten wurden entwickelt und zeigen, dass bei Verwendung eines kurzen Follow-ups Vorhersagen mit hoher Genauigkeit möglich sind. Darüber hinaus zeigen erste Ergebnisse, dass durch einen solchen KI-basierten Klassifikator mehr schnell progrediente Patienten in klinische Studien eingeschlossen werden könnten. Mit diesem Wissen könnten klinische Studien in Zukunft mit deutlich weniger Patienten durchgeführt werden, was zu erheblichen Effizienzsteigerungen und Kosteneinsparungen führen wird.

Es ist das erste Mal in der wissenschaftlichen Literatur, dass eine derart umfassende, multimodale Analyse von Parkinsonverlaufstypen durchgeführt wurde und die Generalisierbarkeit der Ergebnisse an 3 verschiedenen Studien gezeigt werden konnte.

Die Ergebnisse des Projektes wurden auf zahlreichen Kongressen der relevanten Fachgesellschaften (Deutsche Gesellschaft für Neurologie; Deutsche Gesellschaft für Klinische Neurophysiologie; Deutsche Parkinson Gesellschaft; International Movement Disorders Society) in Vorträgen und Postern präsentiert. Die zentralen Ergebnisse wurden international hochrangig publiziert (npj Parkinson's disease) und zusammen mit dem relevanten Quellcode frei zugänglich gemacht. Weitere Publikationen sind aktuell in Arbeit. Konkrete Folgeprojekte mit verschiedenen neuen Kooperationspartnern wurden begonnen und sind in Planung.

Das Projekt wurde in enger Kooperation zwischen der TU Dresden und dem Fraunhofer-Institut SCAI durchgeführt. Der Hauptantragsteller Tom Hähnel (TU Dresden) brachte dabei die klinische Expertise in das Projekt ein und führte die oben beschriebenen Aufgaben im Wesentlichen selbst durch. Das Fraunhofer-Institut SCAI hatte aufgrund seiner datenwissenschaftlichen Expertise eine supervidierende Funktion bei der Vermittlung der datenwissenschaftlichen Methoden an den Hauptantragsteller. Darüber hinaus konnten durch das Fraunhofer-Institut SCAI die entsprechenden Datensätze und Rechenkapazitäten zur Verfügung gestellt werden.

# Eingehende Darstellung des Abschlussberichtes ParKInsonPredict

**Zuwendungsempfänger:** Technische Universität Dresden - Medizinische Fakultät Carl Gustav Carus - Klinik und Poliklinik für Neurologie, Förderkennzeichen: 16DKWN113A

**Vorhabenbezeichnung:** ParKInsonPredict – Entwicklung und Testen eines KI-Modells zur Vorhersage von Krankheitsprogressions-Mustern bei neurodegenerativen Erkrankungen am Beispiel der Parkinsonerkrankung auf Basis von multi-modalen Daten verschiedener Patientenkohorten.

**Laufzeit des Vorhabens / Berichtszeitraum:** 1. September 2022 – 30. Juni 2024

## Aufgabenstellung

Übergeordnete Fragestellung dieses Projektes war, in welchem Maß mit KI-gestützten Algorithmen eine zuverlässige Vorhersage von Krankheitsprogressions-Mustern (z. B. langsam, mittel, schnell bzw. Subtypen je nach im Vordergrund stehenden Beschwerdebild) möglich ist, um die Grundlage für einen zukünftigen Einsatz solcher KI-Modelle in der klinischen Forschung und in der Patientenversorgung zu ermöglichen. Hierzu sollte in einem ersten Schritt aufbauend auf Vorarbeiten des Fraunhofer-Institut für Algorithmen und Wissenschaftliches Rechnen (SCAI) mittels multi-modaler Daten der PPMI-Studie (Parkinson's Progression Markers Initiative) ein KI-Modell entwickelt und evaluiert werden. In einem zweiten Schritt sollte das bereits vortrainierte Modell auf verschiedene Studiendatensätze angewandt werden. Auf diese Art sollte evaluiert werden, in welchem Maße ein solches KI-Modell auf andere klinische Szenarien generalisierbar ist, was eine essentielle Voraussetzung für den zukünftigen Einsatz in der regulären Patientenversorgung darstellt. Ein weiterer Schwerpunkt bestand in der statistischen Untersuchung der Charakteristika der so identifizierten Progressionsmuster, wobei auch hier die Generalisierbarkeit dieser Erkenntnisse über die verschiedenen Studien hinweg evaluiert werden sollte.

Eine solche individuelle Vorhersage der Krankheitsprogression könnte dabei nicht nur für Patienten eine wichtige Information zur Lebensplanung (z. B. berufliche Anpassungen oder Wohnraumanpassungen) darstellen. Auch würde eine solche individuelle Vorhersage von Geschwindigkeit und Art der Symptomprogression die Grundlage zu einer rechtzeitigen Erkennung und zielgerichteten Therapie von neuen oder rasch progredienten Symptomen bei Betroffenen darstellen. Diesem kommt insbesondere im Hinblick auf die hohe Vielfalt der motorischen und nicht-motorischen Parkinson-Symptome eine hohe Relevanz zu. Im Bereich der Forschung würde eine Identifikation von klinischen Verlaufstypen eine gezieltere Selektion von Patienten für zukünftige klinische Studien am Universitätsklinikum Dresden und anderen Kliniken ermöglichen.

Im Hinblick auf aktuell mehr als 200.000 Menschen mit Parkinson in Deutschland<sup>1</sup> und eine erwartete Verdopplung der Zahl von Parkinson-Patienten innerhalb der nächsten beiden Dekaden<sup>2,3</sup> kommt einer Optimierung der medizinischen Versorgung von Parkinsonpatienten eine erhebliche gesellschaftliche Relevanz zu.

## Wissenschaftlicher und technischer Stand, an den angeknüpft wurde

Die Parkinson-Krankheit ist die zweithäufigste neurodegenerative Erkrankung nach der Alzheimerkrankheit.<sup>4</sup> Die Prävalenz wird in Deutschland auf über 200.000 Erkrankte in der Bevölkerung geschätzt und ist stark altersabhängig (65-Jährige: 0,5 %; 75-Jährige: 1-1,5 %; 85-Jährige: 3-4 % der Bevölkerung).<sup>1</sup> Mehrere Studien gehen dabei von einer Verdopplung der Zahl von Menschen mit Parkinson in den nächsten Dekaden aus.<sup>2,5</sup> Im bundesweiten Vergleich wären bestimmte Bundesländer wie z.B. Sachsen mit einem deutlich überhöhten Altersdurchschnitt von 46,9 Jahren hiervon besonders betroffen.<sup>6</sup> Ursache der Parkinson-Krankheit ist primär das Absterben dopaminproduzierender Nervenzellen im Gehirn, welche zu einer über Jahrzehnte chronisch progredient verlaufenden Erkrankung führen. Betroffene entwickeln im Krankheitsverlauf meist eine variable Ausprägung verschiedener motorischer Symptome wie Hypokinesie (Unterbeweglichkeit), Rigor (Muskelsteifigkeit), Ruhetremor (Zittern), posturale Instabilität (Standunsicherheit), Gangstörungen mit der Gefahr von Stürzen, Kamptokormie (Haltungsstörung), Dysphagie (Schluckstörung), Dysarthrie (Sprechstörung), Freezing (plötzliches, kurzes "Einfrieren" bei Bewegungen), Dyskinesien (unwillkürliche, überschießende Bewegungen) und Dystonien (Verkrampfungen/Fehlhaltungen). Eine besondere Herausforderung bei der Erfassung dieser motorischen Symptome besteht in teils erheblichen tageszeitlichen Schwankungen (Fluktuationen), wodurch die während der ärztlichen Konsultation sichtbaren Symptome oft nicht repräsentativ für den Tagesverlauf des Patienten sind. Darüber hinaus können eine Vielzahl weiterer neuropsychiatrischer Symptome mit ebenso variabler Ausprägung bestehen, wie Depressionen, Demenz, psychotische Symptome (Halluzinationen oder Wahnvorstellungen), Apathie (Gleichgültigkeit mit Abwesenheit von Emotionen und Interessen), Angststörungen, Störungen der Impulskontrolle (z.B. erhöhtes sexuelles Verlangen, Esssucht, Spielsucht) und „punding“ (sinnloses, ständiges Wiederholen von Tätigkeiten). Weitere häufige Beschwerden sind eine orthostatische Hypotonie (niedriger Blutdruck im Stehen und ggf. damit verbundene Beschwerden), Harninkontinenz, erektile Dysfunktion, Sialorrhö (verstärkter Speichelfluss mit ggf. Herauslaufen des Speichels aus dem Mund), Obstipation, An-/Hyposmie (Riechstörungen), verschiedene Arten von Schlafstörungen und häufig auch Schmerzstörungen.<sup>7</sup> Eine kurative Behandlungsoption gibt es für die Parkinson-Krankheit aktuell nicht, es existieren aber vielfältige und effektive symptomatische Behandlungsmöglichkeiten für die aufgezählten Symptome.<sup>8</sup> Die komplexen Symptomkonstellationen und Fluktuationen im Tagesverlauf erfordern jedoch eine systematische, zeitaufwendige Erfassung der Beschwerden und darüber hinaus regelhaft die Mitbetreuung durch Spezialisten.<sup>9-11</sup>

Künstliche Intelligenz gewann in den letzten Jahren im Bereich der Medizin und insbesondere im Bereich neurodegenerativer Erkrankungen wie der Parkinson-Erkrankung massiv an Bedeutung und stellt zunehmend den Schwerpunkt vieler aktueller Forschungsprojekte dar. So hat sich die jährliche Zahl der Publikationen zu diesem Thema im Parkinsonbereich seit 2018 nahezu verzehnfacht.<sup>[1]</sup> Entscheidend für diesen Trend sind die zunehmende Verfügbarkeit strukturierter, digital erfasster Behandlungsdaten sowie digitaler Biomarker, welche erst den Einsatz solcher Algorithmen erlauben.

Aktuelle Publikationen fokussieren sich dabei nahezu ausschließlich auf die Erkennung und Graduierung motorischer Symptome mittels verschiedener Sensoren<sup>12,13</sup> und auf diagnostische Fragestellungen wie die Differenzierung zwischen Patienten mit Parkinson-Erkrankung und gesunden Kontrollprobanden<sup>13</sup> sowie die Differenzierung zwischen Patienten mit Parkinson-Erkrankung und Patienten mit atypischen Parkinsonsyndromen<sup>14</sup>. Als Datengrundlage der KI-Modelle dienen dabei neben häufig verwendeten MRT-Bildgebungsdaten<sup>14-16</sup> überwiegend digitale Biomarker, welche mittels Beschleunigungssensoren<sup>17-20</sup> durch

---

[1] Pubmed-Suche nach "Artificial AND Intelligence AND Parkinson" vom 13.11.2021 (pubmed.ncbi.nlm.nih.gov)

Stimmaufzeichnungen<sup>12</sup>, Handschrifterkennung<sup>21</sup> und mittels Kameraaufnahmen<sup>22</sup> erfasst werden. Auch wurden einige KI-Modelle auf Basis von genetischen Informationen trainiert<sup>23,24</sup>.

Durch den Fokus auf digitale Biomarker für motorische Symptome finden dabei ebenso relevante, nicht-motorische Parkinson-Symptome, wie oben ausführlich aufgezählt, wenig Beachtung. Schwerpunkt der genannten Arbeiten ist weiterhin die Diagnosestellung und Differentialdiagnostik. Aus klinischer Sicht kommt nach der Diagnosestellung insbesondere der frühzeitigen Erkennung von neu oder erneut auftretenden motorischen und nicht-motorischen Parkinsonsymptomen eine entscheidende Rolle zu, um frühzeitig eine Therapie einleiten zu können. KI-Modelle zur Vorhersage und Klassifikation solcher klinischen Krankheitsprogressionsmuster auf Basis umfangreicher multi-modaler Daten könnten dazu essentiell beitragen und dabei gleichzeitig eine effiziente Nutzung beschränkter Ressourcen im Gesundheitssystem unterstützen. In diesem Bereich existieren bisher nur wenige Pilotstudien.<sup>25-28</sup> Einige dieser Studien setzen dabei lediglich auf eine Analyse von Baseline-Daten zur Klassifikation der Patienten.<sup>25,26</sup> Nur ein Teil der Studien nutzt direkt longitudinale Daten zur Identifizierung der klinischen Verlaufstypen.<sup>27,28</sup> Diese beiden Studien basierten dabei auf Daten des frei zugänglichen PPMI-Datensatzes, wobei jeweils nur Patienten mit einer neu diagnostizierten Parkinson-Krankheit (Diagnosestellung < 2 Jahre, keine medikamentöse Therapie) in die Analyse eingeschlossen wurden. Die Anwendbarkeit der verwendeten KI-Modelle auf Patienten mit einer bereits länger als 2 Jahre bestehenden Parkinson-Krankheit, welche klinisch die deutlich häufigere Situation darstellt, bleibt dabei unklar. Da beide Studien auf dem gleichen PPMI-Datensatz beruhen, bleibt außerdem unklar, in welchem Maße die entwickelten Algorithmen allgemein auf andere Kohorten mit Parkinsonpatienten generalisierbar sind.

Grundlage dieses Projektes sollte der vom SCAI entwickelte Algorithmus VaDER (Variational Deep Embedding with Recurrence) sein, welcher eine Identifikation und Vorhersage von Krankheitsprogressionsmustern anhand multi-modaler Daten erlaubt und einen speziellen Algorithmus zum Umgang mit unvollständigen Daten implementiert.<sup>27</sup> Die bisherige Arbeit an diesem Algorithmus konzentrierte sich dabei auf methodische Aspekte wie die Genauigkeit der Vorhersage von Clustern bei einem zunehmenden Anteil fehlender Daten. Weiterhin wurde die erfolgreiche Anwendbarkeit auf Parkinson-Daten des PPMI-Datensatzes demonstriert, allerdings ohne dass die so identifizierten Verlaufstypen ausführlich medizinisch charakterisiert wurden. Weiterführende Symptomdomänen wie z.B. kognitive Testergebnisse, Informationen aus MRT-Bildgebungen und nuklearmedizinischen Bildgebungen (DaTSCAN) oder mittels Sensoren erhobene digitale Biomarker wurden nicht weiter berücksichtigt.<sup>27</sup> Im hier beantragten Projekt sollten diese Punkte adressiert werden, indem die identifizierten Verlaufstypen umfassend charakterisiert wurden und die Generalisierbarkeit auf andere Datensätze und Patienten mit fortgeschrittener Parkinson-Krankheit evaluiert wurde.

## Ablauf des Vorhabens

Grundlage des beantragten Projektes war der am Fraunhofer-Institut für Algorithmen und Wissenschaftliches Rechnen (SCAI) entwickelte KI-Algorithmus VaDER (Variational Deep Embedding with Recurrence), der bereits in der Vergangenheit auf dem PPMI-Datensatz zur Identifikation und Vorhersage von klinischen Progressionsmustern trainiert wurde. Im Arbeitspaket 1 wurde dieses Modell auf mehrere weitere Datensätze angewendet, so dass für jeden Patienten eine Zuordnung zu einem klinischen Verlaufstyp erfolgte. In Arbeitspaket 2 wurden die einzelnen Charakteristika (Baseline- und Verlaufskarakteristika) der identifizierten Verlaufstypen mittels statistischer Analysen untersucht und die Robustheit der identifizierten Charakteristika der Verlaufstypen durch einen statistischen Vergleich verschiedenen Studien nachgewiesen. In Arbeitspaket 3 wurde ein Klassifikator zur Vorhersage der Verlaufstypen für die genannten Datensätze

entwickelt und dabei die Generalisierbarkeit eines solchen Klassifikators evaluiert. Arbeitspaket 4 schließlich widmete sich der Veröffentlichung und Verbreitung der Ergebnisse sowie der Planung von Folgeprojekten.

Das Projekt wurde in enger Kooperation zwischen der TU Dresden und dem Fraunhofer-Institut SCAI durchgeführt. Der Hauptantragsteller Tom Hähnel (TU Dresden) brachte dabei die klinische Expertise in das Projekt ein und führte die oben beschriebenen Aufgaben im Wesentlichen selbst durch. Das Fraunhofer-Institut SCAI hatte aufgrund seiner datenwissenschaftlichen Expertise eine supervidierende Funktion bei der Vermittlung der datenwissenschaftlichen Methoden an den Hauptantragsteller. Darüber hinaus konnten durch das Fraunhofer-Institut SCAI die entsprechenden Datensätze und Rechenkapazitäten zur Verfügung gestellt werden.

## Wesentliche Ergebnisse

### **ARBEITSPAKET 1: Transferlernen des existierenden, auf PPMI vortrainierten VaDER Modells**

#### **Task 1.1: Aufbereitung der Datensätze aus DIGIPD und eines Datensatzes des Universitätsklinikum Dresden im für VaDER richtigen Format**

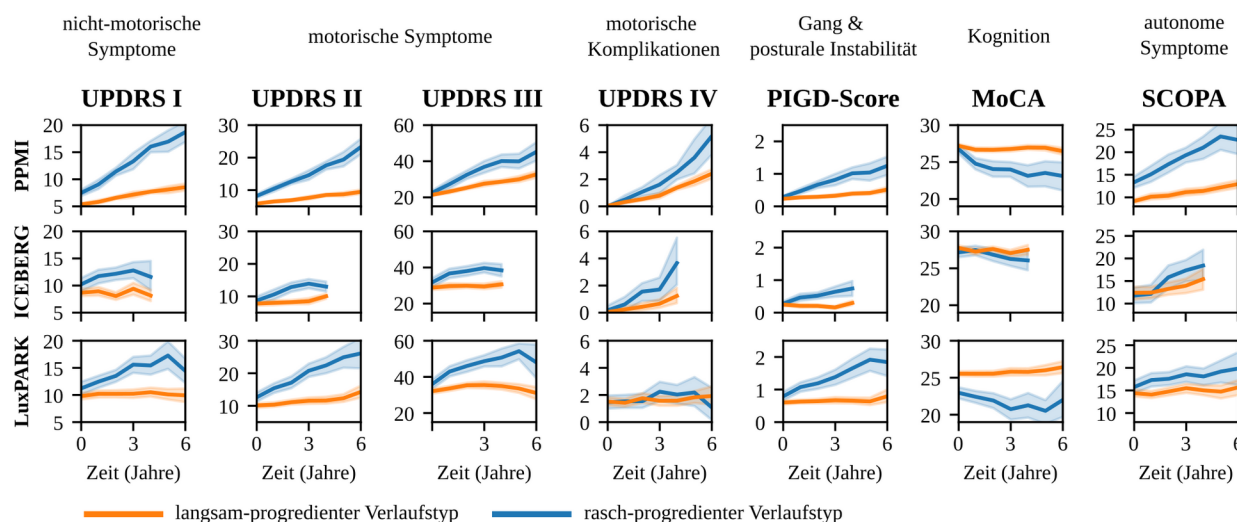
Die Datensätze wurden entsprechend aufbereitet und über alle Studien hinweg homogenisiert. Hierzu wurde beachtet, welche Fragebögen und Tests in welcher der Studien erhoben wurden, sowie die Häufigkeit der vollständigen Erhebung dieser. Entsprechend der Übereinstimmung der erhobenen Daten zwischen den Studien wurden Teil-Scores oder Summenscores nach einheitlichen Vorgehen in allen Studien berechnet, sodass eine Vergleichbarkeit hergestellt wurde. Da sich die Patienten der Studien in unterschiedlichen Krankheitsstadien befanden, wurden diese zunächst mittels eines Latent Time Joint Mixed-effects Model (LTJMM) auf einer einheitlichen Krankheitskala ausgerichtet.

#### **Task 1.2: Transferlernen des VaDER-Modells auf diese Datensätze**

Das VaDER Modell wurde einzeln auf die longitudinalen Daten der entsprechenden Datensätze trainiert (*in-cohort* Ansatz). Dabei wurden jeweils nur die direkt vergleichbaren, d.h. in allen Studien vorhandenen und in Task 1.1 homogenisierten, klinischen Scores verwendet. In jedem der DIGIPD-Datensätze konnte dabei jeweils ein schnell-progredienter und ein langsam-progredienter Parkinson-Verlaufstyp identifiziert werden. Diese unterschieden sich jeweils nur geringfügig in den klinischen Scores zu Beginn der Studie, allerdings erheblich in ihrem Krankheitsverlauf (Abb. 1). Dies zeigt, dass es sich tatsächlich um Verlaufstypen handelt und Patienten nicht nur anhand verschiedener Krankheitsstadien bzw. Baseline-Werte in Typen eingeteilt wurden.

Dadurch, dass in jedem der Datensätze identische Charakteristika und Verlaufstypen festgestellt wurden, ist bereits an dieser Stelle von einer hohen Generalisierbarkeit dieses Verlaufstypen auszugehen. Zur genaueren Evaluation der Generalisierbarkeit, wurde zusätzlich noch ein kohortenübergreifender Ansatz (*cross-cohort*) gewählt. Dabei wurde VaDER auf den frei verfügbaren PPMI-Datensatz trainiert und schließlich auf die anderen Datensätze angewandt. Hierbei zeigten sich weitgehend vergleichbare Ergebnisse wie im *in-cohort* Ansatz.

In dem von der TU Dresden verwendeten Datensatz konnten aufgrund der geringen Patientenzahlen und aufgrund eines hohen Anteils fehlender Werte keine signifikant verschiedenen Verlaufstypen identifiziert werden, sodass dieser Datensatz in den weiteren Analysen nicht weiter betrachtet wurde.



**Abbildung 1: Verlaufscharakteristika der identifizierten Verlaufstypen in den 3 Studien PPMI, ICEBERG und LuxPARK.**

Die Abbildung zeigt den klinischen Verlauf nicht-motorischer Symptome (UPDRS I), vom Patienten berichteter motorischer Symptome (UPDRS II), vom Untersucher erfasster motorischer Symptome (UPDRS III), motorischer Komplikationen (UPDRS IV), Einschränkungen von Gang, Haltung und Gleichgewicht (PIGD), kognitiver Einschränkungen (MoCA) und autonomer/vegetativer Symptome (SCOPA) über die ersten 4-6 Jahre in den Studien PPMI, ICEBERG und LuxPARK.

## ARBEITSPAKET 2: Analyse der Charakteristika der klinischen Verlaufstypen in den einzelnen Studien

### Task 2.1: Charakterisierung der klinischen Verlaufstypen

Um die beiden identifizierten Verlaufstypen detaillierter zu charakterisieren und auch verschiedene motorische und nicht-motorisch Domänen einzubeziehen, wurden auch klinische Fragebögen und Scores betrachtet, die nicht direkt zwischen den Studien vergleichbar waren, also nicht in allen Studien erhoben wurden. Um eine Vergleichbarkeit zu gewähren, wurden diese klinischen Outcomes in 22 verschiedene Symptomdomänen (z.B. Depression, Schlafstörung, Tremor) eingeteilt. Mittels einer Metaanalyse konnten schließlich alle vorher normalisierten Scores einer Symptomdomäne in einen Zahlenwert zusammengefasst werden und schließlich auch zwischen den Studien weiter zusammengefasst werden (Abb. 2). Um auch verschiedene Arten von Fragebögen (einzelne Ja/Nein-Fragen, einzelne Fragen mit abgestuften Antworten, Summenscores von Fragebögen), wurden entsprechend verschiedene mathematische Modelle zur Abbildung der Krankheitsprogression verwendet (linear mixed-effect models, ordinal mixed-effect models, binary mixed-effect models). Unter Nutzung dieser Methodik konnte für die meisten Symptomdomänen ein signifikanter Progressionsunterschied nachgewiesen werden (Abb. 3).

Neben der reinen Analyse klinischer Charakteristika zwischen den Verlaufstypen wurden in einzelnen Studien auch weitere Informationen analysiert: bildgebende Marker („DaTSCAN“), liquorchemische Marker (NfL), digitale Ganganalysen, Sterblichkeit und Therapieansprechen.

Der schnell-progrediente Verlaufstyp wies dabei eine raschere Progredienz in der digitalen Ganganalyse auf. Konkret zeigte dieser Verlaufstyp beim Gehen: geringere Geschwindigkeit, geringeren Zehen-Boden-Abstand, geringeren Fersen-Boden-Abstand, kürzere Schrittweite, flacheren Zehen-Boden-Winkel, längere Standzeit auf dem Standbein, sowie eine kürzere Schwingdauer des anderen Beins. Die Wahrscheinlichkeit des Versterbens war für Patienten des rasch-progredienten Verlaufstyps 3,4x erhöht im Vergleich zum langsamen Verlaufstyp.

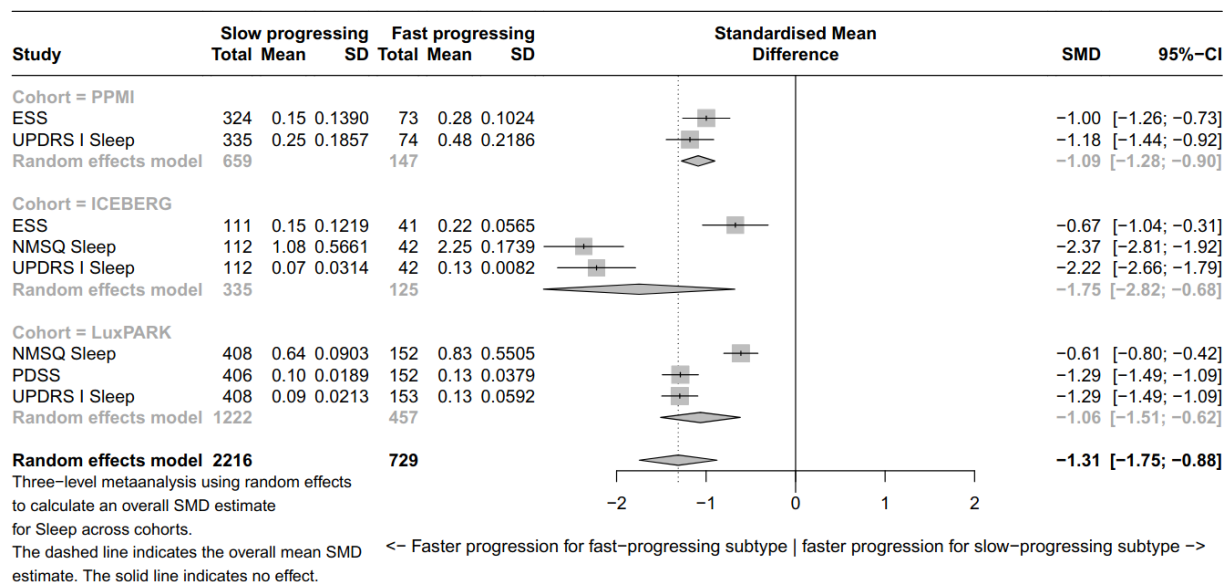
Mittels bildgebender Verfahren (DaTSCAN) konnte eine schnellere Neurodegeneration der Substantia nigra für den rasch-progredienten Verlaufstyp nachgewiesen werden, während der langsam-progredient Verlaufstyp eine starke lateralisierte Neurodegeneration zeigte. Der liquorchemische Marker NfL (Neurofilament light chain) stellt einen allgemeinen, unspezifischen Marker der Geschwindigkeit des Voranschreitens neurodegenerativer Erkrankungen dar. Auch hier bestand ein signifikanter Unterschied mit Zeichen einer rascheren Neurodegeneration des rasch-progredienten Verlaufstyps.

Weiterhin kennzeichnete den rasch-progredienten Verlaufstyp ein schlechteres Ansprechen auf medikamentöse Therapien.

Alle oben beschriebenen Ergebnisse konnten mittels des *cross-cohort* Ansatzes nochmals validiert werden.

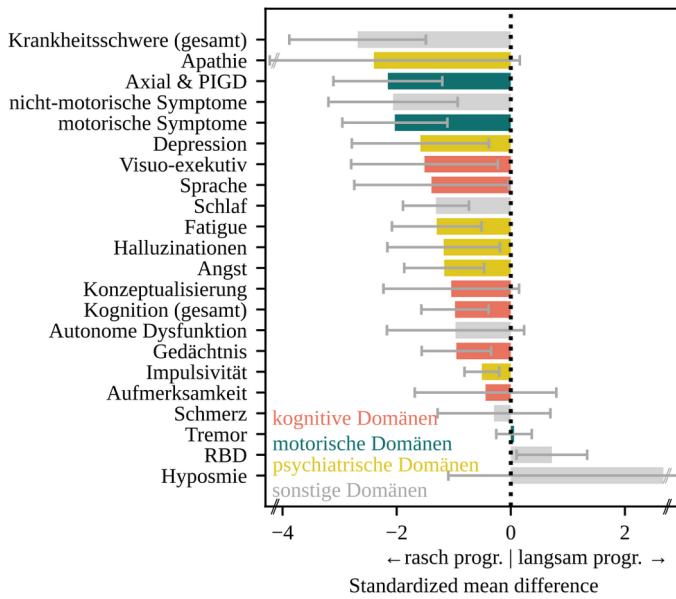
Dies ist das erste Mal, dass in der wissenschaftlichen Literatur eine so umfassende, multimodale Analyse von Parkinson-Verlaufstypen durchgeführt wurde.

Forest plot for progression characteristics of symptom domain Sleep



**Abbildung 2: Beispiel für eine Metaanalyse aller Scores der Symptomdomäne „Schlafstörung“, um die Progression von Schlafstörungen trotz unterschiedlicher klinischer Scores in den Studien vergleichen zu können.**

Innerhalb jeder der Studien wird zuerst aus der Progression einzelner Schlaf-Outcomes ein Mittelwert gebildet. Anschließend wird ein Mittelwert über alle 3 Studien hinweg berechnet, der somit die Informationen aller verfügbaren Scores aller Studien integriert.



**Abbildung 3: Unterschiede der Symptomprogression zwischen dem rasch-progredienten und langsam-progredientem Verlaufstyp**

Ein negativer Wert kennzeichnet dabei, dass der rasche Verlaufstyp eine schnellere Progression bezüglich der dargestellten Symptomdomäne aufweist.

## **Task 2.2: Untersuchung der Generalisierbarkeit der identifizierten Progressionsmuster**

Für motorische und nicht-motorische Progressionsunterschiede der beiden Verlaufstypen konnte in allen 3 Kohorten eine hohe Generalisierbarkeit gezeigt werden. Daten für bildgebende Marker, digitale Ganganalysen, Sterblichkeit und Therapieansprechen lagen jeweils nur in einzelnen Datensätzen vor, sodass hier keine Generalisierbarkeit auf weitere Studien gezeigt werden konnte.

Insgesamt konnte, erstmalig in der wissenschaftlichen Literatur, eine solche umfassende Generalisierbarkeit über verschiedenste Symptombereiche und über mehrere Kohorten hinweg gezeigt werden.

## **ARBEITSPAKET 3: „Training und Evaluation von Klassifikatoren zur Vorhersage der Verlaufstypen“**

### **Task 3.1: Re-Trainieren oder Fine-Tuning eines Klassifikators**

Nach entsprechender umfangreicher Charakterisierung der oben dargestellten Parkinson-Verlaufstypen bestand der zweite Schwerpunkt dieses Projektes in der Evaluation, ob die identifizierten Verlaufstypen auch unter Nutzung von Baseline-Daten oder eines kurzen Beobachtungszeitraums vorhersagbar sind. Hierzu wurden verschiedene konventionelle und *state of the art* Modelle (logistische Regression, Random Forrest, XGBoost) aus dem Bereich des Maschinellen Lernens unter Nutzung einer geschachtelten Kreuzvalidierung auf den vorliegenden Daten trainiert und die Vorhersagefähigkeit dieser Modelle auf den gleichen Datensatz evaluiert (*in-cohort* Ansatz). Um eine Vergleichbarkeit zwischen den Datensätzen zu ermöglichen, wurden jeweils nur die direkt vergleichbaren klinischen Scores verwendet, welche durch ihre hohe Verbreitung auch gut im klinischen Alltag in Zukunft eingesetzt werden könnten.

Zusätzlich zu diesem *in-cohort* Ansatz wurden die Klassifikatoren auch auf den frei verfügbaren PPMI-Daten trainiert und anschließend auf den anderen beiden Studien evaluiert (*cross-cohort* Ansatz).

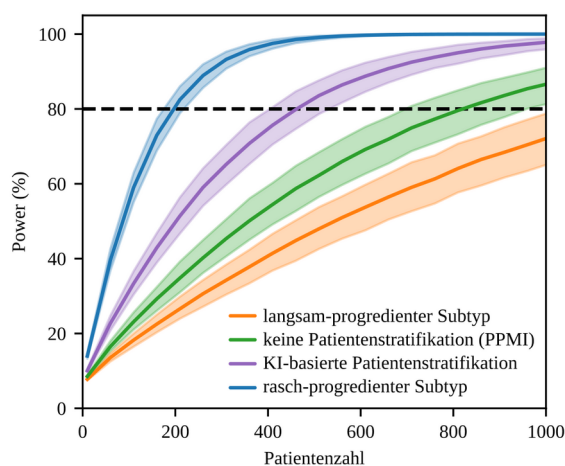
Über alle untersuchten Daten hinweg ergaben sich signifikante klinische Prädiktoren für einen raschen Verlaufstyp: kognitive Einschränkungen (insbesondere visuo-exekutive Funktionseinschränkungen und Sprachdefizite), motorische Einschränkungen (insbesondere sogenannte axiale Symptome) sowie das Vorliegen einer REM-Schlaf-Verhaltensstörung. Weiterhin war eine mittels DaTSCAN nachgewiesene asymmetrische Neurodegeneration hinweisend auf einen raschen Verlaufstyp. Liquorchemische Marker einer zusätzlich zum Parkinson bestehenden Alzheimer-Demenz erhöhten ebenso die Wahrscheinlichkeit für einen raschen Verlaufstyp deutlich. Auch bestimmte genetische Mutationen, wie z.B. die GBA-Mutation, waren mit einem raschen Verlaufstyp assoziiert. Interessanterweise war ein hoher Bildungsstand protektiv, d.h. dieser reduzierte das Risiko für einen raschen Verlaufstyp. Eine weitere umfassende Analyse weiterer Faktoren, inkl. genetischer Risikoscores (polygenic risk scores), verschiedener Komorbiditäten und Medikamente zeigte hingegen keinen Zusammenhang mit einem raschen Verlaufstyp.

### **Task 3.2: Evaluation der Genauigkeit des Klassifikators in den einzelnen Studien**

Entsprechende Vorhersagen der Verlaufstypen waren mit 58% bis 68% Genauigkeit (ROC-AUC) mit Baseline-Werten möglich. Wenn zusätzlich Daten einer weiteren Visite nach einem Jahr eingeschlossen wurden, erhöhte sich diese Genauigkeit auf 67-79%. Im *cross-cohort* Ansatz ergaben sich etwas geringere

Genauigkeiten: 56-61% Genauigkeit bei Verwendung von Baseline-Werten und 70-71% bei Verwendung einer weiteren Visite nach einem Jahr

Während die genannten Genauigkeiten signifikant waren, reichen diese nicht aus, um zuverlässige Aussagen auf Ebene individueller Patienten zu treffen. Daher wurde weiter untersucht, ob eine Nutzung dieser Vorhersagen zur Optimierung klinischer Studiendesigns nutzbar ist. Hierbei zeigte sich, dass durch eine solche Vorhersage Patientenstratifikationen möglich sind, die einen vermehrten Einschluss rasch-progredienter Patienten in klinische Studien ermöglichen. Hierdurch erhöht sich die statistische Power von klinischen Studien und es müssen entsprechend weniger Patienten eingeschlossen werden. In der von uns simulierten Studie ergab sich eine Reduktion der Patientenzahl um 43%. Entsprechend würden sich durch solche Studiendesigns erhebliche Effizienzgewinne und Kosteneinsparungen ergeben.



**Abbildung 4: Optimierung klinischer Studiendesigns durch Einschluss von Patienten mit hoher Wahrscheinlichkeit für einen rasch-progredienten Verlauf.**

Die Abbildung zeigt die statistische Power simulierter klinischer Studien in Abhängigkeit der Patientenzahl (X-Achse) und abhängig von der Art der Patientenrekrutierung (Farbe). Bei Einschluss nur rasch-progredienter Patienten (blau) ergibt sich eine deutlich geringere Patientenzahl als ohne Patientenstratifikation (grün), dies ist aufgrund der nicht hundertprozentigen Genauigkeit der Klassifikatoren jedoch praktisch noch nicht möglich. Allerdings ermöglichen auch die o.g. Klassifikatoren eine deutliche und relevante Reduktion der Patientenzahl (lila). Im Gegensatz hierzu würde eine theoretische Kohorte aus nur langsam-progredienten Patienten eine noch höhere Patientenzahl notwendig machen (orange).

## ARBEITSPAKET 4: „Publikation und Verbreitung der Ergebnisse, Planung von Folgeprojekten“

### Task 4.1: Vorbereitung von Kongressbeiträgen und einer Publikation in einem Fachjournal

Die Ergebnisse des Projektes konnten bereits auf dem DGN-Kongress 2023 (jährlicher Kongress der Deutschen Gesellschaft für Neurologie) in einem Vortrag der „Hot Topics“-Session präsentiert werden. Hierbei ist das breite Publikum des Kongresses, welches auch Nicht-Wissenschaftler und den ärztlichen/wissenschaftlichen Nachwuchs einschließt sowie die hohe Teilnehmerzahl des Kongresses (>7000 Teilnehmende) hervorzuheben. Zielpublikum des Kongresses sind insbesondere alle in Deutschland klinisch tätigen Neurologen, sowie Studierende und Angehörige anderer Fachkreise.

Auf dem Deutschen Parkinsonkongress wurden die Projektergebnisse 2024 gezielt der Parkinson-Fachcommunity vorgestellt. Dies geschah dabei in mehrfachen Formen: als wissenschaftliches Poster, als eigenständiger Vortrag und in Form eines weiterführenden Symposiums zu Digitalen Biomarkern in der Parkinsonforschung.

Im Rahmen weiterer Vorträge wurden die Ergebnisse am Paris Brain Institute, der Universität Luxemburg, dem Deutschen Zentrum für Neurodegenerative Erkrankungen (DZNE), der Universitätsklinik Köln und auf einem Symposium am Fraunhofer SCAI zu Digitalen Biomarkern präsentiert.

Weitere Ergebnisse wurden in Form eines Posters auf dem Internationalen Kongress für Bewegungsstörungen (International Congress of Parkinson's Disease and Movement Disorders) 2023 gezeigt, wobei sie gegenüber einem internationalen Publikum sichtbar gemacht wurden. Eine weitere Posterpräsentation erfolgte auf dem Kongress für Klinische Neurowissenschaften (DGKN Kongress 2024).

Die ParKInsonPredict-Ergebnisse in ihrer Gesamtheit wurden in *npj Parkinson's disease* publiziert, einem der international höchstrangigsten Journale im Parkinson-Forschungsbereich (IF 2023 6,7). Eine weitere Publikation eines Nebenergebnisses zur Diagnoseverzögerung bei Parkinson ist geplant.

Zur Stärkung der Datenkompetenz des wissenschaftlichen Nachwuchses wurden neben o.g. Vorstellungen auf verschiedenen Konferenzen, die auch Nachwuchswissenschaftler adressierten, ein Vortrag im Rahmen des Doktorandenseminars der TU Dresden gehalten und die Arbeit in Form eines Posters zur Summerschool Neurologie der Deutschen Gesellschaft für Neurologie präsentiert.

## **Task 4.2: Aufbereitung des Quellcodes und Erstellung einer Dokumentation zur Veröffentlichung über ein Code-Repository**

Zusammen mit o.g. Publikation wurde relevanter Quellcode zum Training aller Modelle und der statistischen Auswertung publiziert (<https://github.com/SCAI-BIO/PD-progression-types>). Hierdurch soll eine möglichst hohe Transparenz bezüglich des Forschungsprojekts hergestellt werden und gleichzeitig eine Weiterverwendbarkeit durch andere Forschende ermöglicht werden.

## **Task 4.3: Planung von konkreten Folgeprojekten**

Neben unten aufgeführten mittelfristig geplanten Projekten mit verschiedenen Kooperationspartnern wurden auch konkrete Folgeprojekte geplant und teils bereits begonnen. Im Rahmen des ParKInsonPredict-Projektes entstand ein Austausch mit der französischen Hochschule Telecom Sudparis, wodurch ein konkretes Projekt zur Stimmanalyse bei Parkinson-Patienten begonnen wurde. Hierbei liegt das Ziel aufbauend auf existierenden Stimmaufnahmen an beiden Studienzentren in der Entwicklung von Algorithmen zur Vorhersage von Parkinsonsymptomen anhand von Sprachaufnahmen. Ein besonderer Fokus liegt dabei in der Validierung der entwickelten Modelle in der jeweils anderen Sprache (deutsch → französisch, französisch → deutsch).

Nachdem in ParKInsonPredict die Krankheitsprogression bei Parkinson untersucht wurde, soll in einem Folgeprojekt die Progression von Patienten mit REM-Schlaf-Verhaltensstörung (RBD) untersucht werden. RBD stellt eine prodromale Form der Parkinsonkrankheit dar, bei der Betroffene meist innerhalb von 15 Jahren eine Parkinsonkrankheit entwickeln. Auch in diesem Projekt soll die Generalisierbarkeit der gewonnenen Erkenntnisse durch Evaluation in mehreren Kohorten sichergestellt werden. Eine einzelne RBD-Kohorte wurde zum aktuellen Zeitpunkt bereits analysiert, die Analyse weiterer Kohorten stellt nun den nächsten Schritt dar.

Die in ParKInsonPredict gewonnenen Erkenntnisse basieren auf der Untersuchung mehrerer klinischer, standardisierter Studien mit Parkinsonpatienten. Um die Generalisierbarkeit dieser Erkenntnisse in der Regelversorgung zu überprüfen, sollen Daten aus der Parkinsonambulanz der TU Dresden verwendet werden. Hierzu soll auf innovative datenwissenschaftliche Ansätze zurückgegriffen werden, welche die Analyse dieser Daten ermöglichen. Hierzu wurde bisher ein konkreter Projektplan erstellt. Ein Ethikvotum und die notwendigen Daten liegen bereits vor. Ein Antrag für Fördermittel, um im Rahmen dieses Projektes mehrere Doktoranden beschäftigen zu können, wird in den nächsten Monaten gestellt.

## Wichtigste Positionen des zahlenmäßigen Nachweises

- **Personalausgaben:** Die Ausführung von ParKInsonPredict geschah überwiegend durch den Antragssteller Tom Hähnel, was den Großteil der Projektausgaben auf Seite der TU Dresden ausmacht. Dies schließt folgende Aufgaben ein: Datenaufbereitung, Programmierung/Modellierung, statistische Auswertung, Erstellung von Publikationen und Beiträgen für Fachkongresse, Planung von Folgeprojekten. Aufgrund der Laufzeitverlängerung lagen die Personalausgaben leicht über der ursprünglichen Planung, dies konnte allerdings durch Einsparungen bei den Sachmitteln (Verbrauchsmaterial und Reisekosten) ausgeglichen werden.
- **Dienstreisen:** Als zweitgrößter Kostenpunkt waren mehrere Dienstreisen zur Präsentation der Projektergebnisse auf nationalen und internationalen Kongressen notwendig. Um einen optimalen Austausch zwischen den Projektpartnern TU Dresden und Fraunhofer SCAI zu gewährleisten, waren regelmäßige Fahrten zwischen Dresden und Bonn notwendig.
  - International Congress of Parkinson's Disease and Movement Disorders (2023, Kopenhagen)
  - DGN Kongress (2023, Berlin)
  - DGKN Kongress (2024, Frankfurt)
  - Deutscher Parkinsonkongress (2024, Rostock)
  - Vortrag Universität Luxemburg (2024, Luxemburg)
  - Dienstreisen Bonn-Dresden

## Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten

Die im Rahmen des Projekts ParKInsonPredict durchgeführten Arbeiten waren sowohl notwendig als auch angemessen, um die gesteckten Ziele zu erreichen und einen bedeutenden Beitrag zur Parkinson-Forschung zu leisten.

### Notwendigkeit der Projektarbeiten:

- **Bedarf an optimierten Studien zur Erforschung neuer krankheitsmodifizierender Therapien:** Angesichts des Fehlens krankheitsmodifizierender Therapien für die Parkinson-Krankheit und der Vielzahl gescheiterter klinischer Studien in der Medikamentenentwicklung besteht ein dringender Bedarf klinische Studiendesigns zu optimieren. Eine Möglichkeit hierzu stellt eine Patientenstratifikation nach Verlaufstypen dar.

- **Komplexität der Erkrankung:** Die hohe Heterogenität der Parkinson-Symptome und -Verläufe erfordert detaillierte und multidimensionale Analysen. Die Anwendung fortschrittlicher datenwissenschaftlicher Methoden war notwendig, um dieser Komplexität gerecht zu werden und individuelle Krankheitsverläufe besser zu verstehen.
- **Lücken in der bestehenden Forschung:** Vor Beginn von ParKInsonPredict gab es keine solch umfangreiche Analyse von Parkinsonverlaufstypen und keine Studien, die die Generalisierbarkeit von Erkenntnissen zu Parkinsonverlaufstypen über verschiedene Kohorten hinweg untersuchten. Die Notwendigkeit, diese Lücke zu schließen und robuste, generalisierbare Modelle zu entwickeln, machte die geleisteten Arbeiten essenziell.
- **Fehlende datenwissenschaftliche Kompetenz beim wissenschaftlichen Nachwuchs:** Zur Beantwortung der o.g. wissenschaftlichen Fragestellungen sind sowohl tiefgreifende klinisch-neurologische Kenntnisse zur Parkinsonkrankheit als auch tiefgreifende datenwissenschaftliche Kompetenzen notwendig. Letztere existieren beim wissenschaftlichen Nachwuchs in der Parkinsonforschungs-Community, welche sich primär aus behandelnden Ärzten und Ärztinnen zusammensetzt, nicht. Andererseits fehlt es Datenwissenschaftlern (meist (Bio)informatikern, Mathematikern oder Ingenieuren von der Ausbildung her) häufig an Verständnis für medizinische Zusammenhänge. Das ParKInsonPredict-Projekt lieferte hierbei einen entscheidenden Beitrag zum Aufbau einer solchen interdisziplinären Kompetenz auf beiden Seiten. Das Team um Prof. Fröhlich am Fraunhofer Institut SCAI konnte enorm der medizinischen Expertise von Herrn Dr. Hähnel profitieren, während Herr Dr. Hähnel umgekehrt sich wesentlich im datenwissenschaftlichen Bereich weiterbilden konnte.

#### **Angemessenheit der geleisteten Projektarbeiten:**

- **Methodischer Ansatz:** Die Verwendung des VaDER-Algorithmus und die Anwendung auf mehrere unabhängige Datensätze waren angemessen, um die Ziele des Projekts zu erreichen. Die sorgfältige Datenaufbereitung und -harmonisierung sowie Validierung in externen Kohorten gewährleisteten robuste Ergebnisse.
- **Interdisziplinäre Zusammenarbeit:** Die enge Kooperation zwischen der TU Dresden und dem Fraunhofer SCAI kombinierte klinische Expertise mit datenwissenschaftlichem Know-how. Dieser Ansatz war angemessen, um komplexe Fragestellungen effektiv zu bearbeiten und innovative Lösungen zu entwickeln.
- **Umfang der Analysen:** Die umfassende Charakterisierung der Progressionstypen über 22 klinische Symptomdomänen sowie die Einbeziehung von bildgebenden Verfahren, Liquormarkern und digitalen Ganganalysen waren angemessen, um ein tiefgreifendes Verständnis der Krankheitsprogression zu erlangen.
- **Einhaltung von Zeit- und Budgetplänen:** Die Projektarbeiten wurden im vorgesehenen Zeitrahmen und Budget durchgeführt, ohne wesentliche Abweichungen. Dies zeigt die Angemessenheit der Projektplanung und -durchführung.
- **Dissemination und Folgeprojekte:** Die Ergebnisse wurden erfolgreich publiziert und auf Fachkongressen präsentiert. Die Planung und Initiierung von Folgeprojekten unterstreichen die nachhaltige Wirkung und Angemessenheit der geleisteten Arbeiten.

#### **Schlussfolgerung**

Die durchgeführten Projektarbeiten waren notwendig, um wesentliche Fortschritte in der Parkinson-Forschung zu erzielen und die Datenkompetenz des wissenschaftlichen Nachwuchses zu stärken. Die Angemessenheit der Methoden und der interdisziplinäre Ansatz haben maßgeblich dazu beigetragen, die gesetzten Ziele zu erreichen und eine Grundlage für zukünftige Forschungs- und Anwendungsfelder zu schaffen.

## **Voraussichtlicher Nutzen, Verwertbarkeit des Ergebnisses und Verwertungsplans**

### **Künftige klinische Anwendung der Erkenntnisse zum Verlauf der Parkinsonkrankheit an der TU Dresden**

ParKInsonPredict stellt eine im Vergleich zu vorherigen Forschungsarbeiten einzigartig umfassende Untersuchung zu Verlaufstypen der Parkinsonkrankheit dar. Durch den Einschluss von mehr als 1100 Parkinson-Patienten und die Validierung in 3 voneinander unabhängigen Studien wurde eine hohe Robustheit und Generalisierbarkeit der gewonnenen Erkenntnisse gezeigt, welche bisher so noch nicht demonstriert werden konnten. Zusätzlich zur Entwicklung von Modellen zur Vorhersage der identifizierten Parkinson-Verlaufstypen erfolgte eine äußerst umfangreiche Charakterisierung dieser. Diese schloss neben 22 klinischer Symptomdomänen auch liquorchemische Marker, bildgebender Untersuchungen, digitale Ganganalysen, genetische Untersuchungen und externe Faktoren wie Bildungslevel, Komorbiditäten und Vormedikationen ein. Durch den interdisziplinären Schwerpunkt mit klinischer Expertise auf Seite der TU Dresden und datenwissenschaftlicher Expertise auf Seite des Fraunhofer SCAI konnten innovative und neuartige datenwissenschaftliche Methoden (LTJMM, VaDER) genutzt werden um konkrete, klinisch interpretierbare Erkenntnisse zu gewinnen.

Die zuverlässige Prognose des Krankheitsverlaufs bei Parkinson, einschließlich der Vorhersage spezifischer Verlaufstypen, ist für die klinische Behandlung von großer Bedeutung. Die Möglichkeit, den Krankheitsprogress eines Patienten vorherzusagen und diesen mit potenziellen Komplikationen wie Gangstörungen oder kognitiven Beeinträchtigungen in Verbindung zu setzen, ermöglicht es Ärzten, gezielte präventive Strategien zu entwickeln. Dazu könnte die Anpassung der medikamentösen Behandlung gehören, etwa durch die Verringerung der Dopaminagonisten zugunsten von Levodopa. Bei fortgeschrittener Erkrankung könnten sich motorische Symptome und Dyskinesien entwickeln, die eine Sekundärtherapie wie die Tiefe Hirnstimulation oder kontinuierliche Dopaminstimulation mittels subkutaner Geräte erforderlich machen. Entscheidungen für solche Therapien stützen sich vorwiegend auf die Vorhersage des Fortschreitens nicht-motorischer Symptome. Die Identifikation von Patienten mit hohem Risiko für bestimmte Komplikationen wird in Zukunft entscheidend sein, um individuell angepasste Behandlungspläne zu erstellen.

Neben einem tiefgreifenderen Verständnis von Krankheitsprogression bei Parkinson können die gewonnenen Erkenntnisse auch zur Optimierung von klinischen Studiendesigns genutzt werden: In der Parkinson-Forschung besteht ein dringender Bedarf an neuen, krankheitsmodifizierenden Medikamenten. Allerdings machen die breite Vielfalt der Symptome und die individuell sehr unterschiedlichen Krankheitsverläufe die Planung und Durchführung von Phase-II- und Phase-III-Studien besonders schwierig. Die Vielfalt der Symptome erschwert die Festlegung von eindeutigen, univariaten Endpunkten für Studien. Zudem führt die unterschiedliche Symptomatik und deren variabler Verlauf zu einer großen Streuung der Ergebnisse, was dazu führt, dass ein hoher Prozentsatz der Studien (>90%) nicht erfolgreich ist und bisher keine einzige

krankheitsmodifizierende Therapie für Parkinson zugelassen werden konnte. Um diese Herausforderungen zu bewältigen, setzen Pharmaunternehmen verstärkt auf Künstliche Intelligenz (KI). Durch den Einsatz von KI zur Stratifikation von Patienten in homogene Gruppen, wie die durch ParKInsonPredict identifizierten Verlaufstypen, kann die Variabilität und damit die benötigte Größe der Stichproben in klinischen Studien erheblich verringert werden. Sollte der Behandlungseffekt zwischen den Patientensubtypen deutlich variieren, könnten sich zukünftige Behandlungsmethoden auf die Subgruppe mit dem stärksten Ansprechen konzentrieren.

Die klinische Relevanz und Aktualität der Forschungsergebnisse spiegeln sich dabei auch in der o.g. hochrangigen Publikation sowie der Vielzahl von Beiträgen auf Fachkongressen wider.

## **Aufbau datenwissenschaftlicher Forschungskompetenz in der Parkinsonforschung an der TU Dresden und in der Parkinson-Community**

ParKInsonPredict half, wichtige datenwissenschaftliche Forschungskompetenz im Parkinsonforschungsbereich der TU Dresden aufzubauen, insbesondere den Bereich Künstliche Intelligenz, Maschinelles Lernen und Data Science. Es handelt sich hierbei um einen einzigartigen, stark interdisziplinären Forschungsbereich, der sowohl starkes klinisches Wissen im Parkinsonbereich als auch Kenntnisse zu komplexen und innovativen datenwissenschaftlichen Methoden vereint.

Die Stärkung der Datenkompetenz des wissenschaftlichen Nachwuchses beschränkt sich dabei nicht nur auf den Hauptantragsteller Dr. Tom Hähnel, sondern ermöglicht eine gesamte Stärkung der Forschungsgruppe Bewegungsstörungen/neurodegenerative Erkrankungen an der TU Dresden und darüber hinaus. So wurden datenwissenschaftliche Methoden durch Vorträge im Rahmen von Doktorandenseminaren und Präsentationen bei der Summerschool Neurologie der Deutschen Gesellschaft für Neurologie auch anderen Nachwuchswissenschaftlern vermittelt. Im Rahmen eines unten im Bericht weiter ausgeführten Projektes sollen in den nächsten Jahren mehrere Nachwuchswissenschaftler im Bereich des Einsatzes von Künstlicher Intelligenz in der Parkinsonforschung promovieren können und durch den Antragssteller betreut werden. Auch sollen im Rahmen der oben erwähnten Marie-Sklodowska-Curie Graduiertenschule AIPD perspektivisch weitere internationale Doktoranden durch Dr. Hähnel mitbetreut und so die existierende Kooperation mit Fraunhofer SCAI gestärkt werden.

## **Transfer der erlernten Kompetenzen und Methoden auf andere neurodegenerative Erkrankungen und Folgeprojekte an der TU Dresden**

Die im Projekt gewonnene Datenkompetenz der TU Dresden soll Ausgangspunkt für weitere kooperative Folgeprojekte darstellen, da die hierbei erlernten Methoden auf ebenso relevante Fragestellungen und andere Erkrankungen außerhalb der Parkinsonkrankheit anwendbar sind. Konkrete Folgeprojekte liegen dabei im Bereich der Untersuchung von Krankheitsprogression bei anderen neurodegenerativen Erkrankungen, der Untersuchung von Progression bestimmter nicht-motorischer Symptome bei Parkinson, sowie der Untersuchung von Krankheitsprogression in der klinischen Regelversorgung.

So wurde von der TU Dresden bereits ein Förderantrag für ein konkretes Folgeprojekt gestellt, welches die Krankheitsprogression bei Parkinsonpatienten in der Regelversorgung, d.h. außerhalb von standardisierten Studien, untersuchen soll. Im Rahmen dieses Projektes sollen mehrere Nachwuchswissenschaftler an der TU

Dresden promovieren und durch den Antragssteller vor Ort betreut werden. Dem Fraunhofer SCAI würde im Rahmen dieser Kooperation die datenwissenschaftliche Methodik beratend begleiten.

Zusammen mit zwei weiteren Projektpartnern soll noch dieses Jahr ein Antrag für Fördermittel gestellt werden, um in einem nationalen Kooperationsprojekt mit datenwissenschaftlichen Methoden die Krankheitsprogression bei anderen neurodegenerativen Erkrankungen und mit Schwerpunkt auf die Progression bestimmter nicht-motorischer Symptome zu untersuchen. Während die beiden Projektpartner die klinische Expertise im Bereich anderer neurodegenerativer Erkrankungen sowie nicht-motorischer Symptome einbringen, kommt der TU Dresden der datenwissenschaftliche Schwerpunkt zu.

Durch ParKInsonPredict ergab sich eine neue nationale Partnerschaft zwischen TU Dresden und Fraunhofer SCAI, welche die Basis für eine weitere Zusammenarbeit im Bereich neurodegenerativer Erkrankungen bilden wird. Weitere Kontakte wurden zwischen der TU Dresden und der Universität Luxemburg, dem Paris Brain Institute, der Hochschule Telecom Sudparis, dem Universitätsklinikum Köln, dem DZNE-Standort Bonn und dem Universitätsklinikum Köln geknüpft und mehrere konkrete Folgeprojekte im Bereich KI und neurodegenerative Erkrankungen wurden begonnen. Mit Ergebnissen dieser Projekte ist innerhalb der nächsten 1-3 Jahre zu rechnen.

Darüber hinaus ist geplant, die TU Dresden als assoziierten Partner in die oben erwähnte Marie-Sklodowska-Curie Graduiertenschule AIPD aufzunehmen. Dadurch ergäbe sich die Möglichkeit, dass Doktoranden aus dem Bereich der angewandten Informatik direkt Einblicke in den Klinikalltag bekämen und der Antragssteller Dr. Hähnel Anleitung hinsichtlich klinischer Fragen bekommen könnten. Insgesamt ergibt sich also durch ParKInsonPredict eine nationale Partnerschaft zwischen TU Dresden und Fraunhofer SCAI, welche die Basis für eine weitere Zusammenarbeit im Bereich neurodegenerativer Erkrankungen bilden wird.

## **Transfer der entwickelten Methoden auf weitere Bereiche der Medizin**

Während ParKInsonPredict wurden verschiedene datenwissenschaftliche Methoden entwickelt und im medizinischen Kontext validiert. Dies betrifft eine Methodik zum Vergleich der klinischen Charakteristika von Progressionstypen über verschiedene Studien hinweg, für die Situation dass sich die Studienpopulationen in Krankheitsstadien und gemessenen Outcomes signifikant unterscheiden. Weiterhin wurde eine Methodik entwickelt, um Patienten auf einer gemeinsamen Krankheitsskala auszurichten, was die Vorhersage von Krankheitsprogressions-Typen verbessert. Diese genannten Methodiken wurden im Rahmen von ParKInsonPredict zusammen mit dem zur Anwendung benötigten Quellcode frei zugänglich publiziert. Hierdurch werden diese Methodiken der restlichen medizinischen Forschung zugänglich gemacht, wobei sich Anwendungsszenarien in verschiedensten klinischen Bereichen ergeben und der gesamte Forschungsbereich Medizin hiervon profitiert.

## **Während der Durchführung des Vorhabens dem Zuwendungsempfänger bekannt gewordener Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen**

Noch vor Beginn der Durchführung von ParKInsonPredict wurden die Ergebnisse einer multizentrischen Medikamentenstudie zu Prasinezumab publiziert, welche allerdings keinen Effekt auf die Krankheitsprogression zeigen konnten.<sup>29</sup> Zum Ende der ParKInsonPredict-Laufzeit erfolgte eine

Subgruppenanalyse, welche zeigte dass eine Patientenstratifikation nach rasch-progredienten Patienten die Power der Studie erhöht hätte und somit ein positiver Effekt auf die Krankheitsprogression hätte nachgewiesen werden können.<sup>30</sup> Diese Publikation stützt somit nochmals unsere Erkenntnisse, dass eine prospektive Patientenstratifikation zur Optimierung von klinischen Studiendesign essenziell ist.

Weiterhin wurden mehrere Studien zu Progressionstypen der Parkinsonkrankheit publiziert, welche allerdings keine solche umfangreiche Charakterisierung der Progressionstypen durchführten und auch auf weniger Patienten und Kohorten basierten.<sup>31-33</sup> Eine biologische Arbeit gibt Hinweise darauf, dass evtl. verschiedene alpha-Synuklein-Strangvarianten den biologischen Hintergrund der beobachteten Verlaufstypen darstellen könnte, dies muss allerdings in weiteren Studien noch belegt werden.<sup>34</sup>

## Veröffentlichungen

### Publiziert:

- Hähnel, T., Raschka, T., Sapienza, S. et al. Progression subtypes in Parkinson's disease identified by a data-driven multi cohort analysis. *npj Parkinsons Dis.* 10, 95 (2024). <https://doi.org/10.1038/s41531-024-00712-3>

### in Planung (Preprint):

- Hähnel, T., Raschka, T., et al. Predictive Modeling to Uncover Parkinson's Disease Characteristics That Delay Diagnosis. *medRxiv* 2024.03.12.24304165 (2024). <https://doi.org/10.1101/2024.03.12.24304165>

## Literaturverzeichnis

1. Tönges, L., Ehret, R., Lorrain, M., Riederer, P. & Müngersdorf, M. Epidemiologie der Parkinsonerkrankung und aktuelle ambulante Versorgungskonzepte in Deutschland. *Fortschr Neurol Psychiatr* **85**, 329–335 (2017).
2. Dorsey, E. R., Sherer, T., Okun, M. S. & Bloem, B. R. The Emerging Evidence of the Parkinson Pandemic. *JPD* **8**, S3–S8 (2018).
3. Simon, D. K., Tanner, C. M. & Brundin, P. Parkinson Disease Epidemiology, Pathology, Genetics, and Pathophysiology. *Clinics in Geriatric Medicine* **36**, 1–12 (2020).
4. de Lau, L. M. & Breteler, M. M. Epidemiology of Parkinson's disease. *The Lancet Neurology* **5**, 525–535 (2006).
5. Dorsey, E. R. *et al.* Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. *Neurology* **68**, 384–386 (2007).

6. Entwicklung des Median- und Durchschnittsalters im Vergleich. *Bevölkerungsanalysen, Themengebiet: Bevölkerungsstruktur* [https://bevoelkerungsmonitor.sachsen.de/download/12\\_bevoelkerungsentwicklung/statistik-sachsen\\_bevoelkerungsmonitor\\_dsa-vs-median.pdf](https://bevoelkerungsmonitor.sachsen.de/download/12_bevoelkerungsentwicklung/statistik-sachsen_bevoelkerungsmonitor_dsa-vs-median.pdf) (2021).
7. Armstrong, M. J. & Okun, M. S. Diagnosis and Treatment of Parkinson Disease: A Review. *JAMA* **323**, 548 (2020).
8. Jenner, P. Treatment of the later stages of Parkinson's disease – pharmacological approaches now and in the future. *Transl Neurodegener* **4**, 3 (2015).
9. Willis, A. W. *et al.* Neurologist-associated reduction in PD-related hospitalizations and health care expenditures. *Neurology* **79**, 1774–1780 (2012).
10. van der Marck, M. A. *et al.* Effectiveness of multidisciplinary care for Parkinson's disease: A randomized, controlled trial: Multidisciplinary/Specialist Team Care in PD. *Mov Disord.* **28**, 605–611 (2013).
11. Cheng, E. M. *et al.* Association of specialist involvement and quality of care for Parkinson's disease. *Mov Disord.* **22**, 515–522 (2007).
12. Tăuțan, A.-M., Ionescu, B. & Santarnecchi, E. Artificial intelligence in neurodegenerative diseases: A review of available tools with a focus on machine learning techniques. *Artif Intell Med* **117**, 102081 (2021).
13. Belić, M. *et al.* Artificial intelligence for assisting diagnostics and assessment of Parkinson's disease-A review. *Clin Neurol Neurosurg* **184**, 105442 (2019).
14. Xu, J. & Zhang, M. Use of Magnetic Resonance Imaging and Artificial Intelligence in Studies of Diagnosis of Parkinson's Disease. *ACS Chem Neurosci* **10**, 2658–2667 (2019).
15. Vitale, A. *et al.* Artificial intelligence applied to neuroimaging data in Parkinsonian syndromes: Actuality and expectations. *Math Biosci Eng* **18**, 1753–1773 (2021).
16. Zhang, Z., Li, G., Xu, Y. & Tang, X. Application of Artificial Intelligence in the MRI Classification Task of Human Brain Neurological and Psychiatric Diseases: A Scoping Review. *Diagnostics (Basel)* **11**, (2021).
17. Samà, A. *et al.* Estimating bradykinesia severity in Parkinson's disease by analysing gait through a waist-worn sensor. *Computers in Biology and Medicine* **84**, 114–123 (2017).

18. Camps, J. *et al.* Deep learning for freezing of gait detection in Parkinson's disease patients in their homes using a waist-worn inertial measurement unit. *Knowledge-Based Systems* **139**, 119–131 (2018).
19. Rigas, G. *et al.* Assessment of Tremor Activity in the Parkinson's Disease Using a Set of Wearable Sensors. *IEEE Trans. Inform. Technol. Biomed.* **16**, 478–487 (2012).
20. Tsipouras, M. G., Tzallas, A. T., Fotiadis, D. I. & Konitsiotis, S. On automated assessment of Levodopa-induced dyskinesia in Parkinson's disease. *Annu Int Conf IEEE Eng Med Biol Soc* **2011**, 2679–2682 (2011).
21. Drotár, P. *et al.* Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease. *Artificial Intelligence in Medicine* **67**, 39–46 (2016).
22. Sibley, K. G., Girges, C., Hoque, E. & Foltynie, T. Video-Based Analyses of Parkinson's Disease Severity: A Brief Review. *J Parkinsons Dis* **11**, S83–S93 (2021).
23. García-Fonseca, Á., Martín-Jimenez, C., Barreto, G. E., Pachón, A. F. A. & González, J. The Emerging Role of Long Non-Coding RNAs and MicroRNAs in Neurodegenerative Diseases: A Perspective of Machine Learning. *Biomolecules* **11**, (2021).
24. Emon, M. A. *et al.* Clustering of Alzheimer's and Parkinson's disease based on genetic burden of shared molecular mechanisms. *Sci Rep* **10**, 19097 (2020).
25. Fereshtehnejad, S.-M., Zeighami, Y., Dagher, A. & Postuma, R. B. Clinical criteria for subtyping Parkinson's disease: biomarkers and longitudinal progression. *Brain* **140**, 1959–1976 (2017).
26. Tsiouris, K. M., Konitsiotis, S., Koutsouris, D. D. & Fotiadis, D. I. Prognostic factors of Rapid symptoms progression in patients with newly diagnosed parkinson's disease. *Artificial Intelligence in Medicine* **103**, 101807 (2020).
27. de Jong, J. *et al.* Deep learning for clustering of multivariate clinical patient trajectories with missing values. *GigaScience* **8**, giz134 (2019).
28. Zhang, X. *et al.* Data-Driven Subtyping of Parkinson's Disease Using Longitudinal Clinical Records: A Cohort Study. *Sci Rep* **9**, 797 (2019).
29. Pagano, G. *et al.* Trial of Prasinezumab in Early-Stage Parkinson's Disease. *N Engl J Med* **387**, 421–432 (2022).

30. Pagano, G. *et al.* Prasinezumab slows motor progression in rapidly progressing early-stage Parkinson's disease. *Nat Med* **30**, 1096–1103 (2024).
31. Dadu, A. *et al.* Identification and prediction of Parkinson's disease subtypes and progression using machine learning in two cohorts. *npj Parkinsons Dis.* **8**, 172 (2022).
32. Birkenbihl, C. *et al.* Artificial intelligence-based clustering and characterization of Parkinson's disease trajectories. *Sci Rep* **13**, 2897 (2023).
33. Zhou, C. *et al.* Two distinct trajectories of clinical and neurodegeneration events in Parkinson's disease. *npj Parkinsons Dis.* **9**, 111 (2023).
34. Just, M. K. *et al.* Alpha-Synuclein Strain Variability in Body-First and Brain-First Synucleinopathies. *Front. Aging Neurosci.* **14**, 907293 (2022).