Global, Yet Incomplete Overview of Cohort Studies in Parkinson’s disease

Sebastian Heinzel\(^{a,1,*}\), Stefanie Lerche\(^{b,c,1}\), Walter Maetzler\(^{a,c}\) and Daniela Berg\(^{a,c}\)

\(^{a}\)Department of Neurology, Christian-Albrechts-University, Kiel, Germany
\(^{b}\)German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany
\(^{c}\)Department of Neurodegeneration, Hertie Institute for Clinical Brain Research, University of Tuebingen, Tuebingen, Germany

Accepted 9 May 2017

Abstract.

**Background:** Parkinson’s disease (PD) is characterized by heterogeneity and multifactorial longitudinal changes. To identify PD subtypes and factors influencing the disease course, multiple cohort studies have been designed globally. Knowledge about existing cohorts is pivotal to foster collaboration, which may help to advance the understanding of PD.

**Objective:** To raise the awareness about PD cohorts and potential global collaboration opportunities.

**Methods:** Observational cohort studies in clinical PD were identified by a European working group (JPND BioLoC-PD) and through literature search. Using a structured survey investigators of 44 cohorts provided basic information on cohorts and assessments performed.

**Results:** For the 44 cohorts (32% on early/de-novo PD), 14,666 participants (cohorts’ median: 138; range: 23–3,090), a median 1.5-year follow-up interval (0.5–4 years) and a median (planned) observational period of 5 years (1–20 years) were indicated. All studies have assessed motor functions often using rating scales (UPDRS-III; 93% of studies) and less frequently quantitative gait/balance (25%) or fine motor assessments (27%). Cognitive (100%), neuropsychiatric (91%), daily living (78%), sleep (70%), sensory (63%), and gastrointestinal/autonomic (55%) assessments were common and often comparable. Neuroimaging data (82%) and biomaterial (69%) have been collected in many studies. Surprisingly, possible disease modifiers, such as sport/physical activity (11%), have rarely been assessed.

**Conclusions:** Existing data of PD cohorts provide vast collaboration opportunities. We propose to establish a comprehensive, up-to-date, open-access internet platform with easy-to-use search tools of PD cohort descriptions and potentially available data. Bringing researchers together to enable collaborative joint, meta- and replication analyses is timely and necessary to advance PD research ultimately required for an understanding of PD that can be translated into more effective therapies.

Keywords: Parkinson Disease, cohort studies, symptom assessment, biomarkers, information dissemination, social networking, biomedical research

---

**INTRODUCTION**

Parkinson’s disease (PD) is a complex and heterogeneous disease regarding possible etiologies, (neurodegenerative) processes [1] and facets of motor and non-motor symptoms in individual PD patients [2]. Moreover, individual differences in the temporal occurrence and progression of symptoms adds
further heterogeneity to PD [3, 4], which poses statistical challenges to predictive and progression marker research and the stratification of patients into PD subtypes. Importantly, advances in these fields will be a prerequisite for successful disease-modifying, causal and personalized treatment strategies of clinical (and maybe even prodromal) PD as well as the determination of treatment efficacy [5, 6].

Empirical data collected in prospective cohort studies have the potential to reveal findings that fulfill criteria for Class I levels of evidence. Yet, such strong evidence requires studies to be adequately statistically powered. Given the complexity of PD large sample sizes and rigorous replication analyses are required to yield robust findings and meaningful effect sizes [7]. Joint, meta- and validation(replication analyses are the key to true findings [8] that may substantially advance the understanding of PD and may be translated into more effective treatments. Open science and collaboration is integral to this timely and implicitly required attitude and scientific approach in the field of PD cohort studies and in (biomedical) research in general [9].

So far, however, most researchers are often not aware of the many cohort studies with valuable longitudinal data in PD patients that could be a fruitful basis for cooperation. Thus, the present review aims to provide a first, non-exhaustive overview of the globally vivid field of PD cohort studies.

METHODS

We followed two approaches to collect information on observational cohort studies in clinical PD. First, data of 12 European PD cohorts previously collected in 2015 as part of the Joint Programme - Neurodegenerative Disease Research (JPND) BioLoC-PD working group [10] were included in the review. Second, we performed a literature search to identify further cohort studies in PD. Studies/cohorts on the progression of Parkinson’s disease in the clinical phase have been reviewed in 2009 [11]. Since we aimed to now include more recent cohorts, which might not have been included in this previous review, we restricted the search to publications between January 2010 and December 2015. The PubMed search (“longitudinal” AND “Parkinson disease” AND “clinical”) listed 306 publications (310 without the filters “Human” and “Full article”). We restricted our analysis to articles written in English and studies with a published follow-up time of at least one year. We excluded review articles, book chapters, editorials, commentaries, hypothesis papers, meta-analyses and abstract-only publications. Moreover, we excluded studies with a sole cross-sectional design, any treatment or PD management intervention and animal studies. In case a cohort study was published several times, we chose the study with the longest follow-up or if equal, the most recent report. Often multiple reasons for exclusion applied, a meaningful flowchart of the exclusion process is therefore not possible. After abstract screening we selected 109 eligible publications of cohort studies. After full text screening for eligibility and exclusion of already surveyed BioLoC-PD cohorts, we contacted 68 corresponding authors via email once or twice within about 6 months. We described our motivation for this review and attached a structured 1-page survey regarding basic cohort details and assessments performed (multiple choice). Of these 68 cohorts, 28 investigators provided at least one cohort survey resulting in a response rate of 41%. The eligible publications with replies from the corresponding or last authors [3, 12–37] provided survey data of 31 cohorts in clinical PD. Here some principal investigators also provided information of additional cohorts in clinical PD. Moreover, we included the Parkinson’s Progression Markers Initiative (PPMI) cohort details as provided on the study website (http://www.ppmi-info.org/). Thus, in total the present review gives an overview of 44 cohorts in clinical PD and provides details on study characteristics and assessments for each of these cohorts (see Supplementary Table). As based on publications cohorts were categorized as “early/de-novo PD” (at baseline), and cohorts not exclusively investigating early/de-novo PD patients: “spectrum of PD duration”, i.e. range of PD stages or moderate/advanced PD. Studies have been conducted in the US (n = 10), UK (n = 6), Germany (n = 5), The Netherlands (n = 4), Sweden (n = 4), Spain (n = 3), Italy (n = 2), France (n = 2), Canada, China, Japan, Luxemburg, New Zealand, Norway and Singapore (each n = 1), or internationally in a multicenter study (n = 1). The global distribution of these cohorts is illustrated in Fig. 1.

RESULTS

Basic cohort characteristics

The total cumulative sample size of participants included in this review is 14,666 (cohorts’ median: 138 participants, range: 23 to 3,090). It consists
mainly of PD patients, however - not always differentiated - also healthy controls, high-risk individuals, and, in some cohorts, of patients with parkinsonism and neurodegenerative diseases other than PD. Some ongoing studies indicated the planned sample size, but data may not be available for the entire sample, yet. A total sample size of fewer than 100 participants was observed in 39% and ≥100 participants in 61% of cohorts. Sample sizes of the cohorts are illustrated in Fig. 2. Data of the participants have been collected repeatedly with follow-up intervals ranging from 0.5 to 4 years (median: 1.5 years) over a (planned) observational period ranging from 1 year to 20 years (median: 5 years).

**Disease domains and specific assessments performed**

For each of the 44 cohort studies included in the present review, information on specific assessments performed are given in the Supplementary Table. The relative frequencies of (disease) domains assessed and methods used in these cohort studies are shown in Fig. 3.

All of the reviewed cohort studies in clinical PD have assessed motor functions. The severity of motor symptoms has most often been assessed using the most common semi-quantitative rating scale, i.e., (MDS-)UPDRS-III (93% of the studies). Quantitative assessments (gait/balance: 25%, fine motor movements: 27%) have been performed less frequently. Forty-eight percent of the studies used further semi-quantitative rating scales (e.g., Hoehn & Yahr Scale; 23%) and various (quantitative) assessments of motor performances.

Non-motor symptoms have been assessed using detailed and specific or more general assessments of different types (see below). While not indicated...
Fig. 3. Overview of (disease) domains assessed and methods used in 44 longitudinal cohort studies in clinical Parkinson’s disease (PD). We defined ten different domains and present the relative frequency (in percent) of assessments performed in the studies. Further details are provided in the results section and in the Supplementary information. Abbreviations: BAI, Beck Anxiety Inventory; BDI, Beck Depression Inventory; CANTAB, Cambridge Neuropsychological Test Automated Battery; CERAD, Consortium to Establish a Registry for Alzheimer’s Disease neuropsychological battery; CSF, cerebrospinal fluid; FDG-PET, fluorodeoxyglucose positron emission tomography; FP-CIT, dopamine transporter SPECT imaging; GDS, Geriatric Depression Scale; HAAS, Honolulu Asia Aging Study; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; MRI, Magnetic Resonance Imaging; NPI, Neuropsychiatric Inventory; PDQ, PD Questionnaire; PDSS, PD Sleep Scale; PSG, polysomnography; QUIP, Questionnaire for Impulsive-Compulsive Disorders in PD; RBANS, Repeatable Battery for the Assessment of Neuropsychological Status; RBD-SQ, Rapid eye movement Sleep behavior Disorder Screening Questionnaire; SCOPA-AUT, Scales for Outcomes in PD-Autonomic; TCS, Transcranial Sonography; UPDRS, Unified PD Rating Scale; UPSIT, University of Pennsylvania Smell Identification Test.

below the (MDS-)UPDRS-I, which provides semi-quantitative ratings of various non-motor symptoms and non-motor aspects of daily living, has been assessed in 66% of studies.

Neuropsychiatric (depression or anxiety; 91%) assessments have mainly been based on self-report questionnaires, e.g., the Beck Depression Inventory (BDI; 45%). Several other questionnaires have (additionally) been used to assess depressive symptoms and anxiety.

Cognitive testing has been performed in all cohort studies. In addition to tests of global cognition, e.g., Mini-Mental State Examination (MMSE; 75%), studies have often used comprehensive neuropsychological test batteries with three or more tests on specific cognitive functions (80%).

Data on gastrointestinal & autonomic functions have been collected in 55% of studies. Using quantitative measurements and/or self-report questionnaires/ratings orthostatic blood pressure changes have been assessed in 32%, and more general aspects of autonomic dysfunction have been assessed using self-report questionnaires/ratings, such as the SCOPA-AUT (30%) and NMS-Q (12%).

A total of 70% of studies have investigated aspects of sleep (self-report questionnaires) and the Epworth sleepiness scale (41%) has been the most commonly used scale. Symptoms of REM-sleep behavior
disorder (RBD) have also been assessed in a relatively large proportion of the studies. Twenty-five percent have used the RBD screening questionnaire (RBDSQ), and 14% performed polysomnography.

At least some (general) aspects of the sensory system have been assessed in 63% of studies. Particularly olfactory function has been frequently determined quantitatively using Sniffin’ Sticks (25%) and the University of Pennsylvania Smell Identification Test (UPSIT; 27%).

Aspects of daily living (78%) have most often been assessed with the self-report rating scale (MDS-)UPDRS-II (66% of studies). About half of the studies have included a health-related quality of life self-report questionnaire (e.g., PDQ-39 in 30%; EQ-5D in 16%).

(Neuro)Imaging has been performed in 82% of studies with MRI (70%; functional MRI, 27%) and dopamine transporter scans (e.g., FP-CIT, 23%) being the most frequently applied techniques. However, imaging has been performed only in a subset of participants in some cohorts.

Biomaterial has been sampled in 69% of the studies. The most regularly collected material has been blood (61%).

Surprisingly, sport or physical activity in PD patients has very rarely been assessed (11%; quantitative assessments and self-report questionnaires), and methods of assessments varied considerably across studies.

**DISCUSSION**

Through the presentation of this wealth of cohort data on a wide range of aspects in PD we hope to stimulate the research community to increase collaboration and data exchange. If the 44 cohorts of this overview were jointly analyzed, an impressive 5-digit sample size with unprecedented statistical power and various opportunities for stratification and validation/replication could be generated. Here, cohorts with a rather small total sample size of fewer than 100 participants (in 39% of cohorts) might particularly benefit from joint analyses due to increased statistical power. On the basis of collaboration on the existing and comparable longitudinal data in clinical PD as well as on existing precompetitive datasets of pharmaceutical clinical studies [9], many novel and innovative projects can be envisioned.

However, collaborative efforts and interpretation of findings from different studies might be hindered due to lack of comparability between cohort studies and assessments performed. About one third of cohorts focused on early/de-novo PD and its progression, and 68% of cohorts comprised patients with longer disease duration and/or a sample with a range of disease durations. Given the changes in progression characteristics of symptoms and pathologies in the course of PD [38], differences in disease duration/severity between cohorts/patients need to be considered for joint analyses. Here the complex genetic heterogeneity and promoting or protecting factors might play an important role for differences in PD progression between patients.

Some cohorts (so far) only had short observational periods (e.g., only 1 year) or only few follow-up visits (e.g., only 1). Such aspects may additionally lower the statistical power and may affect progression characteristics and comparability of cohorts. Moreover, the participants of the majority of cohorts (89%) were recruited via hospitals/clinics and not from the general or community population and potential biases due to participant recruitment (and retention) should be investigated in collaborative analyses across cohorts.

Regarding assessments this review shows that for most disease domains, e.g., motor (ratings: UPDRS-III), cognitive (testing: MMSE, MoCA) and neuropsychiatric (self-report: BDI/depression), several or even the majority of cohorts could perform joint data analysis as they used the same assessment methods. Further improvement of comparability of studies could be reached by a consensus on a modular set of assessments for longitudinal studies in PD as proposed previously [39]. Moreover, the conversion between different scales (e.g., MMSE-MoCA [40]) could enable comparability of cohorts and create additional opportunities for collaboration. Also, longitudinal changes in assessment measures might often be more comparable between studies than cross-sectional data, as specific methodological and study-specific characteristics within a study should be (rather) constant over time.

For some of the specific disease domains additional aspects should be considered to increase collaboration and to help advancing PD research. Motor assessments are often based on clinical ratings and quantitative assessments could provide more objective measures of motor functions. For analyses across cohorts the lack of compatibility of methods and analytical approaches of quantitative motor measures has to be addressed [41]. However, some of the studies may also have collected quantitative motor data,
e.g., as assessed using sensor- and app-technologies, that may more objectively and precisely quantify PD motor symptoms and their progression than common clinical ratings. The harmonization of these data between cohorts may have particular potential, but the process has to be further developed. For the cognitive domain vast opportunities for collaboration exist as not only testing of global but also of various specific cognitive functions has frequently been performed using comprehensive neuropsychological test batteries. Here, a more detailed overview of specific cognitive tests performed could facilitate collaboration and enable a more differentiated investigation of progressive cognitive deficits and dementia in PD. For assessments of olfactory function, by equating scores of Sniffin’ Sticks and UPSIT performances [42] more than twice as many cohorts could be jointly analyzed compared to only considering identical olfactory test methods. Regarding neuroimaging technical differences (scanners, sequences/protocols) need to be considered for analyses across cohorts. These differences are certainly a challenge, but may be overcome with modern analysis software. Moreover, semi-quantitative assessments, e.g. of medial temporal lobe atrophy [43], using established scales may not need exhaustive imaging harmonization. For biomaterial many aspects of collaboration could be envisaged. For instance, the collaborative transfer of samples to one site for laboratory analysis using one method (e.g. DNA chips for genetic analyses, ELISAs for blood/CSF analyses) may be practical, cost-efficient and increase comparability. Surprisingly few cohorts assessed sport/physical activity and assessments widely differed between studies. We propose (retrospective and prospective) assessments of sport/physical activity to be included in cohort studies in clinical PD. Physical activity is a factor of particular importance since physical activity may ameliorate motor as well as non-motor (e.g. cognition, depression) symptoms [44–48]. Without knowing about the (premorbid) life-style/habits and changes in physical activity due to physiotherapy or other forms of exercise a major source of variance in symptoms or their progression may remain unexplained in PD cohorts. One well-established and validated scale is the physical activity scale for the elderly (PASE) [49], which is also used in large cohort studies including the PPMI study, but also the Framingham heart study. To increase comparability between PD cohort studies, we suggest the use of this semi-quantitative measure. In addition, sensor-based quantitative assessments in everyday situations and during exercise might enable to more accurately and objectively quantify physical activity and its change over time. However, diversity of technologies and analysis algorithms might still hinder analyses across cohorts and methods [50].

While not specifically surveyed further assessments or markers were additionally indicated for some cohorts, e.g. PD family history (3 cohorts), nutrition, and anthropometric measures. Other risk factors including smoking status were not indicated (but possibly assessed in several cohorts) and joint analyses of this factor should also be pursued. Also, further interesting risk factors, such as exposure to pesticides and occupation, were not indicated, their quantification and comparison across cohorts, however, might often be difficult.

Additional biases and limitations

Joint analyses across cohorts may entail the risk of introducing biases and sources of (error-)variance, which may emerge from various different factors and aspects. For instance, differences between cohorts in diagnostic criteria (that may have changed over time [51]) and therapeutic patient-care that may differ between clinics or regions. Recruitment and retention strategies, and selection and motivation of PD patients and healthy controls may differ, e.g. between population-based and hospital/clinical-based cohorts. Practical cohort specific aspects that may affect performances/vigilance and/or motivation, e.g., time of day (morning or (after)noon assessments) and medication intake before/during study visits. Selection biases of patients and healthy controls, biases due to study drop-out and due to observers and laboratory/clinical settings might also have a substantial effect on assessments and may differ between cohorts. Moreover, cultural aspects (including nutrition and other life style factors) and socioeconomic status could be important factors in PD and its progression. The vast majority of cohorts included was from Western countries and we acknowledge that the global representation of this review might be biased. Since ethnic and cultural factors might be important for the understanding of PD progression, their investigation and truly global collaborative research should be encouraged.

A journal publication written in English (regardless of the finding or topic of clinical PD research) was an inclusion criterion of the present review. First, although publications in international journals have increasingly become a worldwide standard, the language-criterion of our search might have biased
this review by favoring Western researchers. Second, some cohorts might have valuable longitudinal data, but rather small sample sizes, and analyses with low statistical power might have remained unpublished. For those cohorts, not journal publications but rather a platform to make the cohorts known and enable collaborative joint data analyses with appropriate statistical power should be encouraged. We emphasize that the present global overview of cohort studies in PD is incomplete. The majority of cohorts identified by the literature search was not included as our survey had a response rate of 41%. Moreover, we acknowledge that other literature search strategies, e.g., also using the term “cohort” instead of “longitudinal” (~3000 publications listed), would have yielded many more PubMed search results and might have led to the identification of additional studies collecting longitudinal data of PD patients.

Since many different partly cohort-specific limitations and biases may apply, we believe it is crucial for analyses across cohorts to not only share the data, but to actively include the researchers and clinicians (who know their cohort best) in a collaborative scientific process from the planning of analyses to the interpretation of results.

In conclusion, the present review provides a global overview of cohort studies in PD. We disclaim being exhaustive with this review in this vibrant field of PD research. This overview as well as the specific cohort details provided is aimed to raise awareness of the plethora of cohorts in PD that already exist and have—at least partly—comparable assessments.

The review is intended to stimulate and foster research collaborations and validation of findings. The opportunities for collaborations based on existing PD cohort data are vast and growing. We therefore propose the design and establishment of a comprehensive, up-to-date, open-access internet platform and database with interactive, easy-to-use search tools of PD cohort information and contact information (without providing actual cohort (raw) data). Such a platform could serve as a means to identify collaboration opportunities and make researchers aware of PD cohorts with comparable assessments. This concept is inspired by the idea that not (raw) data-sharing, but active collaborations, sharing of ideas and (cohort-specific) knowledge, planning/discussing analyses and joining forces with experienced data analysts might best advance PD cohort research. By always involving the original investigators intellectual property might not be an issue, while other legal issues could be overcome by using data-sharing and analysis strategies, such as DataSHIELD (http://www.datashield.ac.uk/). Such a platform where cohort investigators themselves could enter descriptions of cohorts and of available data/assessments might yield a more complete and up-to-date overview than in the present review. Hopefully, PD cohorts not included in this review will in the future also find a representation on such a platform.

Certainly each cohort is in part unique, and methodological differences [9, 10] and other limitations in cohort studies [52] have to be considered when performing the above-mentioned approaches. Yet, due to the complexity and heterogeneity of PD increased statistical power and joint, meta- and validation/replication analyses of large (shared) samples are pivotal and urgently needed for true scientific advances [7, 8]. Thereby, (long-held) beliefs may be challenged and the understanding of PD, PD subtypes and individual differences in the disease progression may be substantially broadened.

Such (ongoing) collaborative efforts should be further encouraged. In the Roman Empire 10 cohorts (each a 300–600-man unit) constituted 1 legion [53]. In this sense, we hope that in the future legions, i.e., the combined (statistical) power of many cohorts, will help to advance our understanding of PD and its progressive nature, the individuality of PD patients and their best possible treatment.

CONFLICT OF INTEREST

The authors have no conflict of interest to report.

FUNDING SOURCES FOR THE STUDY

None.

SUPPLEMENTARY MATERIAL

The supplementary table contains, for each of the 44 cohort studies of the present review, information on basic study characteristics and assessments performed.

The supplementary material is available in the electronic version of this article: http://dx.doi.org/10.3233/JPD-171100.

REFERENCES


A prospective cohort comparison with other phenotypes. *JAMA Neurol.*, **72**, 863-873.


Nat Rev Neurol., 7, 528-534.

J Clin Epidemiol., 46, 153-162.

Mov Disord., 31, 1272-1282.

Mov Disord., 30, 1591-1601.

Front Aging Neurosci., 8, 147.