For archiving in
medRxiv

Prevalence of Late-Stage Parkinson's Disease in the US Health Care System: Insights from TriNetX

Sol De Jesus, MD[1,6*], Annika Daya, MPH,[1*] Liba Blumberger, DrPH, [2] Mechelle M. Lewis, PhD,[1,3,6] Doug Leslie, PhD,[2] Samer D. Tabbal, MD[1,6], Rachel Dokholyan, MPH[1,6], Amanda M. Snyder, PhD, CPC-A[1,6], Richard B. Mailman, PhD,[1,3,6] Xuemei Huang, MD, PhD,[1,3,4,5,6,7]

Departments of Neurology,[1] Public Health Science,[2] Pharmacology,[3] Radiology,[4] Neurosurgery,[5] and the Translational Brain Research Center,[6] Pennsylvania State University-Milton S. Hershey Medical Center, Hershey, PA, USA
Department of Kinesiology,[7] Penn State University, University Park, PA 16802 USA

* These authors contributed equally.

Running Title: Late-stage PD in current healthcare landscape

Corresponding Authors

Sol De Jesus, MD
Department of Neurology
Penn State-Hershey Medical Center
Hershey, PD 17033-0850
Tel: 717-531-0003, ext. 287082
Email: sdejesus@pennstatehealth.psu.edu

Xuemei Huang, M.D., Ph.D.
Department of Neurology
Penn State University, H037
500 University Drive, Hershey, PA 17033-0850
Phone: 717-531-0003, ext. 287082
Email: xuemei@psu.edu

Word Count:

	Current	Journal
Title characters	97	100 (characters)
Running head characters	40	45
Key words	4	5
Abstract (words)	249	250
Text (words)	3,137	3,700
Tables & Figures (number)	4 (2 figures, 2 tables)	5
References	50	No limits

Key words: Parkinson's disease, late-stage Parkinson's disease, health care burden, palliative care

Financial disclosure/conflict of interest: Sol De Jesus has received fees as an educational consultant from Medtronic Inc. Xuemei Huang has received research funding from the NIH relevant to this study (U01 NS112008). Richard Mailman was a consultant for D_1 agonist-related matters for Cerevel Therapeutics, the latter company just acquired by AbbVie, Inc.

Funding Sources: None

Disclaimer: Dr. Lewis now is employed by the National Institute of General Medical Sciences (NIGMS). Her role in this study was as an employee of Penn State University. The opinions expressed in this article are those of the authors, and do not reflect the views of the National Institutes of Health or any other external organization.

Abstract

Background: Patients in late-stage Parkinson's disease (PD_{LS}) are caregiver dependent, have low quality of life, and higher health care costs.

Objective: To estimate the prevalence of PD_{LS} patients (per Coelho and Ferreira 2012) in the current United States (US) health care system.

Methods: We downloaded the 2010-2022 data from the TriNetX Diamond claims network that consists of 92 USA health care sites. PD was identified using standard diagnosis codes, and PD_{LS} was identified by the usage of wheelchair dependence, personal care assistance and/or presence of diagnoses of dementia. Age of PD_{LS} identification, and survival information are obtained and stratified by demographic and the disability subgroups.

Results: We identified 1,031,377 PD patients in the TriNetX database. Of these, 18.8% fit our definition of PD_{LS} (n=194,297), and 10.2% met two or more late-stage criteria. Among all PD_{LS}, the mean age of PD_{LS} identification was 78.1 (±7.7), and 49% were already reported as deceased. PD_{LS} patients were predominantly male (58.5%), with similar distribution across PD_{LS} subgroups. The majority did not have race (71%) or ethnicity (69%) information, but for the available information, >90% (n=53,162) were white, 8.2% (n=5,121) Hispanic/Latino, 7.8% (n=4,557) black, and <0.01% (n=408) Asian. Of the PD_{LS} cohort, 71.6% identified with dementia, 12.9% had personal care assistance, and 4.8% were wheelchair bound.

Conclusions: Late-stage patients are a significant part of PD landscape in the current US healthcare system, and largely missed by traditional motor-based disability staging. It is imperative to include this population as a clinical, social, and research priority.

Introduction

Parkinson's disease (PD) is characterized clinically by motor dysfunction that includes bradykinesia, tremor, and rigidity. The 20th century witnessed tremendous advances in understanding the role of nigrostriatal system in PD,[1-4] as well as developing therapeutic strategies targeting at normalizing it.[5-11] The first breakthrough was the availability of levodopa that effectively restored the lost dopaminergic tone.[2] This had *led to a new era for PD patients* with dramatically improved symptoms and lengthened lifespan.[5-11] As of the first quarter of 21st century, PD patients remained largely independent in the first decade of their disease, especially when treated in PD subspecialty clinics and research centers.[12]

In clinical practice, however, PD patients suffer from increasing motor (e.g., motor fluctuations, levodopa-induced dyskinesia, gait disorders, and falls)[13] and non-motor (e.g., cognitive, mood, and autonomic) signs and symptoms.[14,15] Together, this impacts overall independence and quality of life (QoL).[14,15] It is estimated that PD affects at least six-million people worldwide,[16] a number that may double globally by 2040.[16] In a recent GBA report, the increasing PD prevalence has outpaced all other neurological disorders, akin to a pandemic.[17] The increase has been attributed to several factors including an aging population, improved symptom detection, decline in smoking prevalence,[18] and a variety of environmental perturbations,[19-22] although much remains to be learned.[17,18]

Prior to the use of levodopa (the standard of care for more than 50 years), PD was diagnosed based on clinical signs and symptoms. The relentless progression of PD was tracked with Hoehn and Yahr (H&Y) staging that focuses on limited motor and postural instability milestones.[5] With the widespread use of levodopa, definition of PD also involved clinical drug response as well as pathological presence of α-synuclein (**αSyn**)-positive Lewy bodies/neurites and dopamine neuron loss in the substantia nigra pars compacta (**SNc**) of the BG.[3,4] Despite recent awareness of extra-nigrostriatal involvement and non-motor disabilities, PD staging has still largely focused on HY staging. This was the motivation for a recent Movement Disorders Society-sponsored meeting that addressed the many issues involved in developing a more precise PD staging system.[23]

Coelho and Ferreira[15] more recently proposed a very useful alternate terminology for PD staging. They labeled early-stage PD (PD_{ES}) as beginning with initial diagnosis that is marked by substantial motor improvement when patients take levodopa; this is often called the "honeymoon"

period because most patients have excellent QoL. Patients then progress to advanced-stage PD (PD_{AS}) that is marked by levodopa-related motor complications (e.g., motor fluctuations and drug-induced dyskinesia). For this stage of patients, deep-brain stimulation (DBS) attenuates symptoms and can improve QoL dramatically. Finally**,** patients enter late-stage PD (PD_{LS}) marked by levodopa-unresponsive postural instability, dysphagia, prominent autonomic dysfunction and/or dementia, and the need for assistance in walking or even being bed-bound.

The late stage of disease (i.e., PD_{LS}) has received less attention in clinical research, partly due to the complexity of clinical symptoms. Late-stage disability may emerge as early as 7-10 years after diagnosis,[14,15,24,25] and leads to PD_{LS} patients having increased cost and greater utilization of healthcare systems.[26] Analysis of Medicare claims-based data has provided a picture of the financial impact of advanced/late PD (incurring greater mean costs) as compared to mild or moderate PD patients.[27] In addition to falls and dementia, PD_{LS} often is characterized by dependence on caregivers for activities of daily living. Wheelchair dependence and institutionalization are also common. These disability milestones have been suggested to precede death in PD patients by ~5 years.[15,28] The exact prevalence of patients in later stages is, however, unknown, thus hindering effective need assessments for clinical, research, and social planning.

Leveraging the available TriNetX Diamond Network data, we have estimated PD_{LS} prevalence in the U.S health care system. We identified all PD patients, and from these data, determined those in late stages based on the disability milestones of wheelchair dependence, dementia, and/or personal care assistance. We then identified the characteristics of the PD_{LS} patient population to analyze their profiles.

Methods

We utilized the TriNetX Diamond Network, a de-identified clinical patient dataset from a network of healthcare organizations in the US.[29] This database provides ambulatory and primary care electronic medical records, medical claims, and patient medication data from 92 healthcare organizations in the US and includes data for 212-million patients. According to the TriNetX website, the Diamond Network contains data covering 1.8 million providers and 99% of US health plans. We extracted data on February 24, 2023, and then identified and described individuals with PD.

A preliminary query was performed on the TriNetX website to look at the overall PD cohort using the ICD-10 and ICD-9 codes for PD (G20; 332.0). The presence of PD_{LS} was identified via TriNetX query tool with the following combinations of ICD-10 and ICD-9 (Table 1) diagnostic codes: a diagnostic code for PD (G20; 332.0) in combination with a code for wheelchair dependence (Z99.3; V46.3), dementia (F02.80; F02.81; 294.11; 294.10), and/or personal care assistance (Z74.1; Z74.2; Z74.3; V60.89; V60.4). Because PD diagnosis before age 40 is rare (often with atypical features and progression), those individuals were excluded although this may have eliminated some patients with young onset PD. For transparency and better understanding of source data, we also captured all data, even if incomplete or with otherwise inaccurate data entry (Figure 1).

The resulting data library for PD_{LS} patients was requested for the following additional analyses. First, individual PD_{LS} subgroups, based on our querying criteria, were extracted without overlap (PD + one late-stage criterion). Second, we created combined PD subgroups to capture the presence of multiple late-stage criteria (multiple criteria groups) encompassing all possible combinations: (a) dementia and wheelchair; (b) wheelchair and personal care; (c) dementia and personal care; and (d) presence of all three criteria. These data are presented in a Venn Diagram (Figure 2).

Descriptive statistics were used to summarize sample characteristics with frequencies computed for demographic and death data that were stratified by late-stage criteria (dementia, wheelchair dependence, personal care assistance, and their combinations). Binary variables were created to identify whether an individual was alive or deceased at the time of data retrieval. Age was calculated by subtracting a patient's year of birth from their last service date documenting an interaction with the healthcare system partnering with TriNetX. All descriptive analyses were conducted using SAS Version 9.4, and final data was drawn February 24, 2023. Penn State Institutional Review Board prior approval was unnecessary due to the de-identified nature of the database.

Results

Identification of the PD_{LS} cohort and disability distribution in TriNetX dataset

From the preliminary TriNetX data query, 1,031,377 patients between 2010 and 2022 had a diagnosis of PD. Within the overall PD group, 777,116 subjects did not have any PD_{LS} features. After excluding 42,524 subjects who were < 40 yo, 211,737 PD_{LS} patients > 40 yo remained. We also excluded another 17,440 patients: these include patients with an encounter day(s) documented after the reported date of death (N = 11,805); those with only a G80, but no other late-stage code (N = 5,487); and those who died before age 40 (N = 148). This gave a final dataset of 194,297 PD_{LS} patients (Figure 1).

Of the total PD_{LS} cohort, more than 70% (139,023) were assessed as having dementia, 9,301 subjects were identified as having wheelchair dependence, 139,023 subjects were identified with a diagnosis of dementia, and 25,145 subjects were specified utilization of personal care assistance (Figure 2). Among PD_{LS} patients, 10.2% (n=19,787) had at least two of these other late-stage criteria (Figure 2). Within those with multiple diagnoses, 3.1% (n=5,995) had both the diagnosis of dementia and wheelchair dependence; 0.7% (n=1,318) were wheelchair dependence and required personal care assistance; 6.4% (n=12,474) had dementia and personal care assistance, and 0.5% (n=1,041) met all three criteria (Figure 2).

Age and vital status of PD_{LS} cohort and its disability subgroups in TriNetX dataset

The mean age was 77.9 ± 7.7 at the time of PD_{LS} identification in TriNetX database (Table 2). The mean age was slightly lower for those having wheelchair dependence (74.7 ± 9.5 y) and personal care assistance (74.7 ± 7.1 y). The mean age was higher in those diagnosed with dementia (78.9 ± 7.1 y) and those with multiple diagnoses (78.1 ± 7.6 y).

By February 2023, 45% percent (n=88,071) of all PD_{LS} were reported to be deceased. About half of PD_{LS} patients with dementia had died at the time of analysis (49.2%, n=68,368) followed closely by those with multiple disabilities needs/diagnoses (43%, n=9,027). Those with only wheelchair dependence (36.2%, n=3,369) or personal care assistance (29.1%, n=7,307) had fewer deaths (Table 2). For those who had died, the average age at death was 81.8 ± 6.6 y, similar across all groups although slightly lower in the Wheelchair and Personal Assistance groups (Table 2)

Demographic characteristics of PD_{LS} cohort and disability subgroups in TriNetX dataset

A majority of the PD_{LS} patients were male [58.4% (n=113,406) versus 41.6% female (n=80,812) (Table 2)]. The pattern of male predominance seemed to persist in all PD_{LS} subgroups, except for those with PD_{LS} + wheelchair dependence that were sex equivalent. Race was unknown for 71.9% (n=139,630) of PD_{LS} patients; 25.8% (n=50,035) were documented as White, 2.2% (n=4,236) as Black or African American, and 0.2% (n=396) were Asian. Similarly, ethnicity was reported sparsely for PD_{LS} patients: there were no reports for 69.9% (n=135,812); 27.6% (n=53,635) reported as not Hispanic or Latino, and only 2.5% (n=4,850) Hispanic or Latino. [Statistical analyses are in Supplemental tables]

Discussion

Lewis et al. (this issue[12]) highlight the current landscape of PD progression in the early 21st-century, especially noting that PD patients remain largely independent in the first decade of disease. There is, however, a paucity of data and clinical metrics capturing the real-world PD progression in patients beyond 10 years, and/or in later stages of disease. In the current study, we have used the TriNetX clinical/claims database to capture and characterize patients in late-stages of disease (PD_{LS}), using the staging system advocated by Coelho and Ferreira.[15]

Our findings suggest a significant proportion of PD patients (~20%) meet criteria for PD_{LS}. This is consistent with more limited studies both within[27,30] and outside the US.[26] As we discussed below, the variability in the past approach may have led to an under-estimation of PD_{LS} in the past studies. Given that the number of PD patients is expected to double globally by 2040, there likely will be dramatic increases in both the prevalence and healthcare cost that occur for people with PD_{LS}.[26] Thus, it is imperative to include late stage PD population as clinical, social and research priority as a potential and emerging health emergency.

PD_{LS} as a new clinical stage entity: guide the assessment of clinical and society needs.

With improved pharmacologic treatment, advanced therapies, and increasing access to care, patients with PD are living longer and reaching later stages of the disease. Accurate diagnosis and disease staging, as the disease evolves in real world with these new treatments, improved our understanding of health care and research resource allocation. In 2012, Coelho and Ferreira (C&F)[15] advocated for a new staging system to define early, advance and late-stage of PD patients.

Although both sensible and simple to use, the C&F staging system has not been widely adopted, and H&Y staging is still the prevailing PD staging instrument in most if not all clinical trials.

Our current data underscores the limitation of a motor-centric staging, as embodied by H&Y staging, by demonstrating its inadequacy to capture greater than 95% in late stages with unmet needs. Our data also pointed out that dementia is composed of the largest group (>75%) of PD patients in the late stage. Recent studies clearly suggested that those with PD-normal cognitive function may have a normal life expectancy, whereas those with PD-MCI or PD-dementia may have reduced life expectancy.[31] Together, our findings support the clinical and research priorities of dementia in PD, converging with the current NIH priority of Alzheimer's disease and related disorders.

As the research on better staging terminology is still underway for PD patient,[23] clinical coding practices vary significantly among providers. There is a lack of standardization as well as non-specific coding for PD to capture disease-associated manifestations. The adoption of H&Y staging in mid-1990s reflects both the biological progression and the effects of available therapeutic approaches focused on motor disability. The current data advocates the adoption of C&F-like staging system, for its simplicity as H&Y and its ability to capture the multidimensional disability in the new era, in clinical coding system.

In 2021, Dashtipour et al.[32] highlighted the limitations that exist with the current ICD-10-CM coding for PD, particularly in delineating motor complications with disease progression. Inclusion of disability milestones separately from the standard PD diagnostic code (G20) can be an initial way to capture more accurately the patient status and disease severity. With improved capability of tracking larger populations afforded by the expansion and accessibility of claims-based databases (e.g., TriNetX), accurate coding becomes integral to understand the impact of interventions and management and informing research on PD subpopulations like PD_{LS}.

According to a recent global health (GBA) report, PD is the fast growth neurological disorder, outpaced all other neurological disorders, including ADRD.[33] PD_{LS} patients are anticipated to have increased both prevalence and health care system utilization, highlighting the urgent need for improved identification, and tracking of this proportion of PD patients in the current health care system. The results will help guide more equitable, and impactful health care resource allocation.

PD_{LS} as a biomedical opportunity: an understudies population in current research paradigm.

In 2021, the MDS committee report noted that "[s]taging is based on key points within the natural history of a disease, allowing the prediction of future relevant disease milestones … [and] can also be pivotal in developing therapeutic approaches as it can partition individuals into groups with similar outcomes, and it allows for quick and effective communication with patients about disease severity and expected clinical outcomes including, in some instances, treatment responsivity."[23] There is clearly a major unmet need for a modern staging system that allows more reliable assessment of diagnosis and progression. Current research efforts focus primarily on early-stage PD. The new data added the question of how many patients are in the late stage of disease and what is their fate in current research paradigm.

As detailed in our companion paper,[12] the treatment options for symptom relief, disease-modifying therapies, and neuroprotective interventions have continued to evolve.[34] The existing approaches, however, are not focused on PD_{LS} patients. PD_{LS} patients may forego specialized care under the assumption there are limited care options and available pharmacologic treatments.[14] Consistent with the mindset, care for PD_{LS} patients shifts to providing support and comfort, often with poor QoL. This propels the needs for multidisciplinary individualized treatment goals and identifying appropriate pharmacologic and nonpharmacologic interventions carefully chosen from existing toolsets for late stage PD patients.[35]

In addition to being limited to supportive or palliative care options, PD_{LS} patients are usually excluded from pharmaceutical trials. Of the available treatments, most are pharmacological and targeted at dementia (e.g., rivastigmine, donepezil, galantamine), drug-induced psychosis (e.g., clozapine, pimavanserin), depression, and/or urinary/gastrointestinal dysfunction,[13] and orthostatic hypotension (e.g., fludrocortisone, midodrine, droxidopa) with various levels of success. Both PD_{AS} and PD_{LS} are known to involve specific motor and nonmotor symptoms, as well as functional consequences such as repeated falls, increased dependence, mild dementia, and neuropsychiatric symptoms.[20,36] The varied non-motor symptoms (e.g., psychiatric, autonomic) experienced in $PD_{AS/LS}$ impact heavily on both patient and caregiver burden, and are a strong predictor of need for institutional admission.[36,37]

Current research has yet to address the needs of PD_{LS} patients who represent nearly a quarter of the current PD population, estimated to double by 2040.[2] Deep brain stimulation is an

established surgical treatment option including for patients over the age of 75 years with 13 years of illness.[38,39] PD_{LS} patients, however, experience disability PD_{LS} milestones, and are excluded from, or considered ineligible, for DBS-focused PD research.[40,41] In clinical practice, many PD_{AS} and PD_{LS} patients with/without DBS and their caregivers still look for treatment and research opportunities.[13,42] Although, potential disease-modifying therapies and neuroprotective interventions in early and even earlier (prodromal stages)[43,44] continue to be sought, it is unclear when and whether they will address the unmet needs for patients who progress to PD_{LS}.

Limitations and future directions

We are cognizant of several limitations of our analyses. Inherent to real world claims-based datasets, significant demographic information (e.g., race, ethnicity) were under-reported. Indeed, there was no race and ethnicity information for two-thirds of the subjects, making it impossible to determine if a specific underserved population could be identified.[45] In addition, inaccuracies exist in clinical diagnosis and data entry, and there are limitations in the diagnostic codes for PD and PD-associated symptoms within electronic medical record (EMR) systems.[11,46,47] We also could only estimate the true date of diagnosis as it is possible that this was often earlier than the date entered in the TriNetX database if the provider had not been part of the Diamond Network. Additionally, certain diagnostic codes may be under-utilized, contributing to under-estimates of disease progression, although the driving factors for such practices are unclear and should be further studied to improve EMR data capture.

Other cognitive disability milestones such as psychosis/hallucinations were not included in this analysis because heterogeneity and limitations inherent in the definition and coding practices likely are amplified in such databases.[48] These data were obtained from multiple health care organizations in the US, but may not fully represent the overall PD population since a third of the US population was not captured by this EMR database. Therefore, the number of PD_{LS} estimated based on ICD codes in our study is certainly an under-estimate. Relevant to this point, the current diagnostic criteria in the databases are likely to have other parkinsonisms lumped into the diagnosis of PD. Unfortunately, only a limited number of patients have a diagnosis that was confirmed pathologically or via current sophisticated biomarkers. Finally, our data also suggest that there may be either unequal access to healthcare, or markedly under-reported prevalence of PD, in several racial or ethnic groups, an issue that requires attention.

Conclusions

Our analyses show that the prevalence of PDLS is significant, and thus its impact will increase markedly in the future. Future clinical and research priorities should include the PD_{LS} to improve understanding of healthcare utilization patterns, medication status, cost-burden, and intervention-based outcomes. These investments will be meaningful in clinical setting as we can choose the current most effective therapies, either pharmacological, surgical, and supportive, for PD_{ES}, PD_{AS} and PD_{LS}, respectively. Scientifically, we should embrace the challenges of growing PD_{LS} presence and capture potential of emerging novel dopaminergic agents[13,42,49,50] and devise future strategies to address their unmet needs. The implementation of new diagnostic codes (e.g., like early, advanced, and late-stages as proposed by Coelho and Ferreira[15]) may provide simpler and defined disability milestones that encourage tracking of disease progression, inform health economics burden, and identify supportive treatment options in real world health care system.

References

1. Erro R, Picillo M, Vitale C, et al. The non-motor side of the honeymoon period of Parkinson's disease and its relationship with quality of life: a 4-year longitudinal study. *Eur J Neurol* 2016;23(11):1673-1679.
2. Rascol O, Payoux P, Ory F, Ferreira JJ, Brefel-Courbon C, Montastruc JL. Limitations of current Parkinson's disease therapy. *Ann Neurol* 2003;53 Suppl 3:S3-12; discussion S12-15.
3. Calne DB, Snow BJ, Lee C. Criteria for diagnosing Parkinson's disease. *Ann Neurol* 1992;32 Suppl:S125-S127.
4. Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry* 1992;55(3):181-184.
5. Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. *Neurology* 1967;17(5):427-442.
6. Marras C, McDermott MP, Rochon PA, et al. Survival in Parkinson disease: thirteen-year follow-up of the DATATOP cohort. *Neurology* 2005;64(1):87-93.
7. Sweet RD, McDowell FH. Five years' treatment of Parkinson's disease with levodopa. Therapeutic results and survival of 100 patients. *Annals of internal medicine* 1975;83(4):456-463.
8. Barbeau A. High-level levodopa therapy in Parkinson's disease: five years later. *Trans Am Neurol Assoc* 1974;99:160-163.
9. Barbeau A. Six years of high-level levodopa therapy in severely akinetic parkinsonian patients. *Arch Neurol* 1976;33(5):333-338.
10. De Pablo-Fernandez E, Tur C, Revesz T, Lees AJ, Holton JL, Warner TT. Association of Autonomic Dysfunction With Disease Progression and Survival in Parkinson Disease. *JAMA Neurol* 2017;74(8):970-976.
11. Hughes AJ, Daniel SE, Blankson S, Lees AJ. A clinicopathologic study of 100 cases of Parkinson's disease. *Arch Neurol* 1993;50(2):140-148.
12. Lewis MM, Cheng XV, Du G, et al. Clinical progression of Parkinson’s disease in the early 21st century: Insights from AMP-PD dataset. *Movement Disorders* 2023;*(Under review)*.
13. Lewis MM, Van Scoy LJ, De Jesus S, et al. Dopamine D(1) Agonists: First Potential Treatment for Late-Stage Parkinson's Disease. *Biomolecules* 2023;13(5).
14. Schrag A, Hommel A, Lorenzl S, et al. The late stage of Parkinson's -results of a large multinational study on motor and non-motor complications. *Parkinsonism Relat Disord* 2020;75:91-96.
15. Coelho M, Ferreira JJ. Late-stage Parkinson disease. *Nat Rev Neurol* 2012;8(8):435-442.
16. Collaborators GBDN. Global, regional, and national burden of neurological disorders, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* 2019;18(5):459-480.
17. Bloem BR, Okun MS, Klein C. Parkinson's disease. *Lancet* 2021;397(10291):2284-2303.

18. Rossi A, Berger K, Chen H, Leslie D, Mailman RB, Huang X. Projection of the prevalence of Parkinson's disease in the coming decades: Revisited. *Mov Disord* 2018;33(1):156-159.

19. Siderowf A, Concha-Marambio L, Lafontant DE, et al. Assessment of heterogeneity among participants in the Parkinson's Progression Markers Initiative cohort using alpha-synuclein seed amplification: a cross-sectional study. *Lancet Neurol* 2023;22(5):407-417.

20. Tanner CM. Advances in environmental epidemiology. *Mov Disord* 2010;25 Suppl 1:S58-S62.

21. Kasten M, Chade A, Tanner CM. Epidemiology of Parkinson's disease. *Handb Clin Neurol* 2007;83:129-151.

22. Tanner CM, Aston DA. Epidemiology of Parkinson's disease and akinetic syndromes. *Curr Opin Neurol* 2000;13(4):427-430.

23. Cardoso F, Goetz CG, Mestre TA, et al. A Statement of the MDS on Biological Definition, Staging, and Classification of Parkinson's Disease. *Mov Disord* 2023.

24. Reinoso G, Allen JC, Jr., Au WL, Seah SH, Tay KY, Tan LC. Clinical evolution of Parkinson's disease and prognostic factors affecting motor progression: 9-year follow-up study. *Eur J Neurol* 2014.

25. Evans JR, Mason SL, Williams-Gray CH, et al. The natural history of treated Parkinson's disease in an incident, community based cohort. *J Neurol Neurosurg Psychiatry* 2011;82(10):1112-1118.

26. Martinez-Castrillo JC, Martinez-Martin P, Burgos A, et al. Prevalence of Advanced Parkinson's Disease in Patients Treated in the Hospitals of the Spanish National Healthcare System: The PARADISE Study. *Brain Sci* 2021;11(12).

27. Dahodwala N, Li P, Jahnke J, et al. Burden of Parkinson's Disease by Severity: Health Care Costs in the U.S. Medicare Population. *Mov Disord* 2021;36(1):133-142.

28. Fabbri M, Kauppila LA, Ferreira JJ, Rascol O. Challenges and Perspectives in the Management of Late-Stage Parkinson's Disease. *Journal of Parkinson's disease* 2020;10(s1):S75-S83.

29. TriNetX. Search of TriNetX Network of electronic medical records. TriNetX, LLC is compliant with HIPAA and ISMS-only de-identified patient records were sued, and study was exempted from Institutional Review Board approval. January 2022. 2022.

30. Dahodwala N, Pettit AR, Jahnke J, et al. Use of a medication-based algorithm to identify advanced Parkinson's disease in administrative claims data: Associations with claims-based indicators of disease severity. *Clin Park Relat Disord* 2020;3:100046.

31. Backstrom D, Granasen G, Domellof ME, et al. Early predictors of mortality in parkinsonism and Parkinson disease: A population-based study. *Neurology* 2018;91(22):e2045-e2056.

32. Dashtipour K, Bahroo LB, Pagan FL, Torres-Yaghi Y, Gupta F, Isaacson SH. Parkinson Disease ICD-10-CM Coding. *Practical neurology* 2021;September 2021:40-48.

33. Calabrese VP. Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. *Neurology* 2007;69(2):223-224; author reply 224.

34. McFarthing K, Rafaloff G, Baptista M, et al. Parkinson's Disease Drug Therapies in the Clinical Trial Pipeline: 2022 Update. *Journal of Parkinson's disease* 2022;12(4):1073-1082.

35. Hwang H, Norris SA. Managing Advanced Parkinson Disease. *Journal of geriatric psychiatry and neurology* 2021;34(4):289-300.

36. Williams L, Qiu J, Waller S, Tsui D, Griffith J, Fung VSC. Challenges in managing late-stage Parkinson's disease: Practical approaches and pitfalls. *Aust J Gen Pract* 2022;51(10):778-785.

37. Marras C, Beck JC, Bower JH, et al. Prevalence of Parkinson's disease across North America. *NPJ Parkinsons Dis* 2018;4:21.

38. Mitchell KT, Younce JR, Norris SA, et al. Bilateral Subthalamic Nucleus Deep Brain Stimulation in Elderly Patients With Parkinson Disease: A Case-Control Study. *Oper Neurosurg (Hagerstown)* 2020;19(3):234-240.

39. Schuepbach WM, Rau J, Knudsen K, et al. Neurostimulation for Parkinson's disease with early motor complications. *N Engl J Med* 2013;368(7):610-622.

40. Aquino C, Luo L, Agarwal P, et al. Deep Brain Stimulation as an Exclusion Criterion in Parkinson's Disease Studies: Time to Rethink? *Mov Disord* 2023;38(9):1602-1605.

41. Mahlknecht P, Foltynie T, Limousin P, Poewe W. How Does Deep Brain Stimulation Change the Course of Parkinson's Disease? *Mov Disord* 2022;37(8):1581-1592.

42. Huang X, Lewis MM, Van Scoy LJ, et al. The D1/D5 Dopamine Partial Agonist PF-06412562 in Advanced-Stage Parkinson's Disease: A Feasibility Study. *Journal of Parkinson's disease* 2020;10(4):1515-1527.

43. Prasad EM, Hung SY. Current Therapies in Clinical Trials of Parkinson's Disease: A 2021 Update. *Pharmaceuticals (Basel)* 2021;14(8).

44. Gouda NA, Elkamhawy A, Cho J. Emerging Therapeutic Strategies for Parkinson's Disease and Future Prospects: A 2021 Update. *Biomedicines* 2022;10(2).

45. Klinger EV, Carlini SV, Gonzalez I, et al. Accuracy of race, ethnicity, and language preference in an electronic health record. *J Gen Intern Med* 2015;30(6):719-723.

46. Hill EJ, Sharma J, Wissel B, et al. Parkinson's disease diagnosis codes are insufficiently accurate for electronic health record research and differ by race. *Parkinsonism Relat Disord* 2023;114:105764.

47. Rajput AH, Rozdilsky B, Rajput A. Accuracy of clinical diagnosis in parkinsonism--a prospective study. *Can J Neurol Sci* 1991;18(3):275-278.

48. Aarsland D, Batzu L, Halliday GM, et al. Parkinson disease-associated cognitive impairment. *Nat Rev Dis Primers* 2021;7(1):47.

49. Mailman R, Huang X, Nichols DE. Parkinson's disease and D1 dopamine receptors. *Curr Opin Investig Drugs* 2001;2(11):1582-1591.

50. Mailman RB, Yang Y, Huang X. D1, not D2, dopamine receptor activation dramatically improves MPTP-induced parkinsonism unresponsive to levodopa. *European journal of pharmacology* 2021;892:173760.

Figures and legends

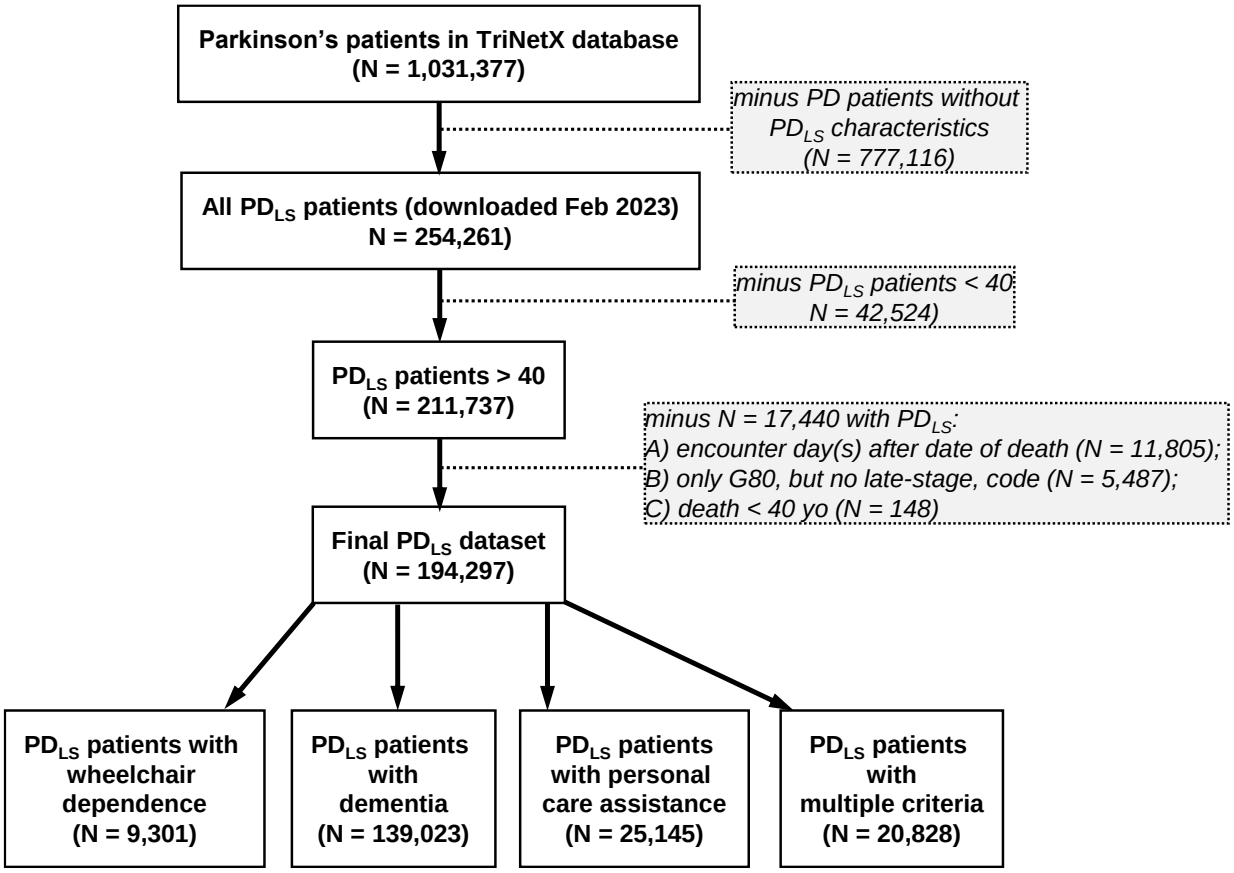
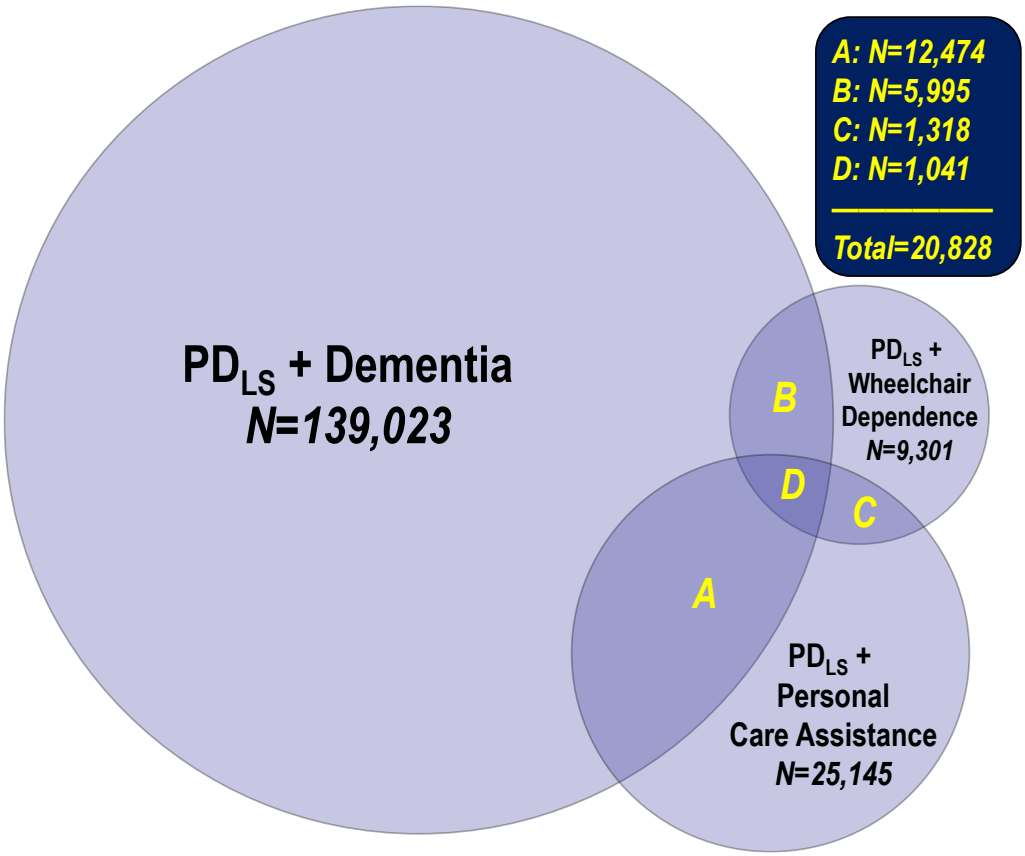


Figure 1: PRISMA Flow Diagram of Cohort Exclusions, Data from the TriNetX database. Data were downloaded and included information from patients with PD_{LS}. Since this study focused on the late-stage characteristics of patients with Parkinson's disease, we excluded patients without PD_{LS} characteristics, those aged <40, and those where the reported date of death occurred prior to a documented encounter. Patients then were separated by the PD_{LS} characteristics that included patients with wheelchair dependence, dementia, personal care assistance, or a combination of these (see Figure 2).



<u>Figure 2. Venn diagram describing overlap in diagnoses and late-stage disease criteria among the PDLS patients.</u>

Patient numbers in each overlapped are inclusive only of those with the listed combination. That is, section A (12,474 patients in PD_{LS}+Dementia vs. PDLS+Personal Care Assistance) does not include N= 1,041 who also have wheelchair dependence.

Tables

Table 1: ICD-10 and ICD-9 codes used in the query parameters of the TriNetX Diamond dataset*.

	ICD-10 Codes	ICD-9 Codes Conversation
Parkinson's disease	G20 Parkinson's disease	332.0. Parkinson's disease
Personal Care Assistance	Z74.1. Need for assistant with personal care	V60.89. Other specified housing or economic circumstances
	Z74.2. Need for assistant at home and no other household member able to render care	V60.4. No other household member able to render care
	Z74.3. Need for continuous supervision	V60.89. Other specified housing or economic circumstances
Wheelchair dependence	Z99.3. Dependence on wheelchair	V46.3. Wheelchair Dependence
Dementia	F02.80. Dementia in other disease classified elsewhere, unspecified severity, without behavioral disturbance, psychotic disturbance, mood disturbance, and anxiety	294.10. Dementia in conditions classified elsewhere without behavioral disturbance
	F02.81. Dementia in other diseases classified elsewhere, unspecified severity, with behavioral disturbance	294.11. Dementia in conditions classified elsewhere with behavioral disturbances

**Note, this data has been accessed via the online database on February 24th, 2023*

Table 2: Demographic and vitals information of subjects identified as late-stage Parkinson's disease (PD_{LS}) in the Diamond TriNetX database (2010-2022).

	Late-stage PD		With Dementia		With Wheelchair Dependence		With Personal care Assistance		With Multiple Criteria		P-value
N	194,297		139,023		9,301		25,145		20,828		
Sex											
Female	80,812	*41.6%*	55,766	*40.1%*	4,686	*50.4%*	11,314	*45.0%*	9,046	*43.4%*	<0.001*
Male	113,406	*58.4%*	83,201	*59.9%*	4,614	*49.6%*	13,819	*55.0%*	11,772	*56.5%*	<0.001*
Unknown	79	*0.0%*	56	*0.0%*	1	*0.0%*	12	*0.1%*	10	*0.1%*	0.74
Race											
Asian	396	*0.2%*	303	*0.2%*	20	*0.2%*	38	*0.2%*	35	*0.2%*	0.02*
Black or	4,236	*2.2%*	2,547	*1.8%*	257	*2.8%*	748	*3.0%*	684	*3.3%*	<0.001*
White	50,035	*25.8%*	34,514	*24.8%*	2,451	*26.4%*	7,473	*29.7%*	5,597	*26.9%*	<0.001*
Unknown	139,630	*71.9%*	101,659	*73.1%*	6,573	*70.7%*	16,886	*67.2%*	14,512	*69.7%*	<0.001*
Ethnicity											
Hispanic or Latino	4,850	*2.5%*	3,072	*2.2%*	321	*3.5%*	752	*3.0%*	705	*3.4%*	<0.001*
Not Hispanic or Latino	53,635	*27.6%*	36,745	*26.4%*	2,738	*29.4%*	8,084	*32.2%*	6,068	*29.1%*	<0.001*
Unknown	135,812	*69.9%*	99,206	*71.4%*	6,242	*67.1%*	16,309	*64.9%*	14,055	*67.5%*	<0.001*
Age (y) of PD_{LS}	77.9 ± 7.7		78.9 ± 7.1		74.7 ± 9.5		74.7 ± 9.3		78.1 ± 7.6		<0.001**
Vital Status											
Alive	106,226	*54.7%*	70,655	*50.8%*	5,932	*63.8%*	17,838	*71.0%*	11,801	*56.7%*	<0.001*
Deceased	88,071	*45.3%*	68,368	*49.2%*	3,369	*36.2%*	7,307	*29.1%*	9,027	*43.3%*	<0.001*
Age at Death (± SD)	81.2 ± 6.6		81.4 ± 6.3		79.1 ± 8.4		79.7 ± 8.2		81.2 ± 6.7		<0.001**

The data represent the number (frequency) and proportion (percent) of patients with noted characteristic(s).
Multiple criteria = PD_{LS} with one or more of dementia, personal care assistance, and/or wheelchair dependance.
Abbreviations: Dx = diagnosis, PD = Parkinson's disease, y = year, **= data unavailable.
Data for age and age at death reflect the mean ± standard deviation (SD).