

Review

Will Artificial Intelligence Replace the Movement Disorders Specialist for Diagnosing and Managing Parkinson's Disease?

Matt Landers^{a,*}, Suchi Saria^{b,c,d} and Alberto J. Espay^{e,*}

^a*Department of Computer Science, Whiting School of Engineering, Johns Hopkins University, Baltimore, MD, USA*

^b*Departments of Computer Science and Statistics, Whiting School of Engineering, Johns Hopkins University, Baltimore, MD, USA*

^c*Department of Health Policy and Management, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA*

^d*Bayesian Health, New York, NY, USA*

^e*Department of Neurology, James J. and Joan A. Gardner Family Center for Parkinson's Disease and Movement Disorders, University of Cincinnati, Cincinnati, OH, USA*

Accepted 13 June 2021

Pre-press 28 June 2021

Abstract. The use of artificial intelligence (AI) to help diagnose and manage disease is of increasing interest to researchers and clinicians. Volumes of health data are generated from smartphones and ubiquitous inexpensive sensors. By using these data, AI can offer otherwise unobtainable insights about disease burden and patient status in a free-living environment. Moreover, from clinical datasets AI can improve patient symptom monitoring and global epidemiologic efforts. While these applications are exciting, it is necessary to examine both the utility and limitations of these novel analytic methods. The most promising uses of AI remain aspirational. For example, defining the molecular subtypes of Parkinson's disease will be assisted by future applications of AI to relevant datasets. This will allow clinicians to match patients to molecular therapies and will thus help launch precision medicine. Until AI proves its potential in pushing the frontier of precision medicine, its utility will primarily remain in individualized monitoring, complementing but not replacing movement disorders specialists.

Keywords: Artificial intelligence, big data, machine learning, Parkinson's disease

*Correspondence to: Matthew Landers, Department of Computer Science, Whiting School of Engineering, Johns Hopkins University, Baltimore, MD, USA. Tel.: +1 201 417 6700; E-mail: mlander5@jhu.edu. and Dr. Alberto J. Espay, Department of Neurology, Gardner Family Center for Parkinson's Disease and Movement Disorders, University of Cincinnati, Cincinnati, OH, USA. Tel.: +1 513 558 4035; E-mail: alberto.espay@uc.edu.

Artificial intelligence (AI) algorithms continue to proliferate in neurological research and health care. In fact, AI-based approaches have emerged in innumerable facets of healthcare including clinical decision support [1], disease detection from imaging [2], and the reduction of disparities in care [3]. One can thus be forgiven for wondering if human and machine sources of "intelligence" are destined to clash and whether

AI might emerge as the more reliable and accurate source of diagnoses. By extension, it is natural to consider whether AI might relegate movement disorders specialists to a secondary role in caring for patients with Parkinson's disease (PD). Arbitrating the relationship between AI and clinicians must resolve three conceptual nodes: 1) big data, 2) validation, and 3) the meaning of a diagnosis.

BIG DATA

AI's central utility relies on its ability to process big data into logical pieces of information that can be interpreted into previously accepted clinical labels. Importantly, the data on which AI-based algorithms operate must be more representative of a patient's disease burden than those accrued during a clinical visit. Furthermore, the data must be an accurate source of the construct of interest. For example, the data continuously captured and stored by portable smartphones is often transformed into a measure of activity. The validity of big data is often assumed given the proliferation and wide usage of smartphones from which data is derived. However, assuming a measure of *activity* can be distilled from these data is dangerous if researchers do not account for the many potential variables spuriously affecting the reliable use of a smartphone during action and inaction.

Consequently, we must recognize that *big* data does not necessarily equal *good* data. The translation of big data into digestible information may be inaccurate or even misleading if not done with high quality methods. Nevertheless, AI-based measures of activity may be helpful in assisting epidemiologic efforts by screening for parkinsonism at a population level and in remote monitoring of already-diagnosed patients [4]. In and of themselves, at present, the data cannot usefully be used to suggest, let alone confirm or refute a diagnosis—at least not with the sophistication required at a clinical level [5].

VALIDATION

If AI-based tools are to be used as endpoints in clinical trials, contribute to regulatory approvals, identify at risk populations, and inform the allocation of medical resources, validation of an algorithm's performance is of critical importance. However, validating any AI output is particularly difficult for a disease with a progression characterized by long time horizons. Moreover, the gold standard

clinician-based measure of PD is subjective and rater-dependent [6, 7]. Thus, researchers must develop methodologies for validating AI-based decisions especially when they are not consistent with in-clinic measurements (a scenario that should be expected if AI is to supplant clinicians in diagnosing molecular disease subtypes). This is a problem for which there is no clear resolution. The validation strategies used for clinical scales and questionnaires, where a new instrument must correlate with another already-validated measure, may not be sufficient for the validation of AI-generated outputs. While recognizing that the use of a "biomarker" identified from a population might not be appropriate for application to an individual [8], researchers may still need to compare AI outputs with those of established techniques such as quantitative MRI/SPECT [9], corneal confocal microscopy [10], camera tremor magnification [11], and retinal nerve fiber layer thickness [12].

DIAGNOSIS

The diagnosis of PD remains a clinical judgment. This judgment is based on a neurological examination at the bedside to positively ascertain supportive clinical features and judiciously rule out exclusionary clinical features [5]. A positive PD diagnosis cannot be more certain than probable, even with the use of a DAT SPECT as ancillary testing, as dopaminergic deficiency is shared with other parkinsonian disorders beyond PD. Thus, clinicians must accept increasing variability in motor and nonmotor presentations, which neither directly aligned with biological markers nor accurately predict response to treatment [13]. Consequently, biomarker validation (usually of peptides, such as alpha-synuclein, amyloid-beta 42, and tau) anchored on clinical diagnosis has yielded poor reproducibility within and between cohorts [14]. Further complicating the diagnosis of PD is the body of evidence that demonstrates pure alpha-synuclein pathology is the exception rather than the rule, with a median of 3 (out of up to 9) pathologies in patients with autopsy-confirmed PD [15]. Moreover, criteria meeting Alzheimer's disease pathology is present in nearly 80% of those with autopsy-confirmed PD [16], thus blurring the diagnostic distinction between PD and Alzheimer's disease. While many AI models have been developed to measure PD based on motor symptoms and other phenotypic evidence [17], they cannot allay these fundamental limitations of diagnosis. Until AI is sufficiently developed as to elucidate

the range of molecular pathways in PD, it can only assist in the movement disorder specialist's clinical judgement. AI models have been demonstrably useful in detecting changes in PD symptoms [18], but the diagnosis of PD, for now, remains clinical rather than biological.

WHAT IS THE BEST PATHWAY FOR AI IN MANAGING AND DIAGNOSING PD?

Treating PD as a single disease has been helpful for the development of symptomatic therapies, which target common denominators, most often dopamine deficiency, but it has been futile in disease modification efforts as unique biological targets may be pathogenic in some, but not in most of those affected. Hence, it is important to separate the role of AI in assisting *symptomatic* versus *disease-modifying* efforts.

Symptom monitoring and management - Yes

Fundamentally, AI algorithms can take a vast array of input data and classify patients based on the features relevant to a therapy's effect. In this way, these techniques minimize the reliance on clinician judgement and speculation for characterizing and modeling a patient's disease burden.

Traditional symptom monitoring efforts have relied on in-clinic patient evaluation. While clinical assessments are important for treating patients with PD, the measures collected during these visits can only serve as a proxy for a patient's daily experience. An authentic examination of disease burden requires the context of a patient's typical day. By continuously collecting movement, voice, and other relevant data [19], smartphones, and the sensors embedded therein, generate a vast amount of data that permit an accurate and objective assessment of a patient's movement impairment [20], speech limitation [21], somnolence [22], and other disease-related burdens. Because AI-based tools are designed to draw conclusions from high dimensional data, they have been effective in using sensor data to "measure" disease. For example, algorithms have been designed to quantify Parkinson's disease motor symptom severity [23].

The management of a disease requires an accurate and thorough understanding of a disease's effects on the way a patient feels and functions. It is thus not surprising that the Food and Drug Administration has continued to emphasize the necessity to capture a patient's experience in the development of therapies

[24, 25]. Before AI, efforts to understand an individual's disease burden required methods that were often able to detect larger changes in symptom severity but insufficient for subtler shifts in a patient's quality of life [26]. The introduction of AI-based methods has allowed for pattern-detection unencumbered by commonly accepted anchors (e.g., "ON medication state", "wearing off", or "diphasic dyskinesia") that may more appropriately represent a patient's experience [23].

Disease understanding and disease modification - Not yet

Approaches for disease modification can only succeed by targeting pathogenic molecular/biological pathways. These pathways, however, are rarely common across populations defined solely by clinical traits. PD is not a single disease but a collection of diseases, thus disease modification demands a match between molecular therapies and the relevant disease-causing biology, even beyond the underlying genotypes (e.g., LRRK2-PD is associated with substantial phenotypic diversity [27]).

The reduction of several molecularly distinct disorders into a single disease is a significant limitation in realizing the dream of precision medicine for PD and beyond [28]. A central principle of this approach to disease management is the stratification of diseases into subtypes. With statistical and AI-based methods [29, 30] researchers have identified molecular subtypes in diseases such as pancreatic ductal adenocarcinoma, a condition for which subtypes are particularly important for predicting a patient's response to chemotherapy [31].

The identification of disease subtypes is a promising frontier. In the near term, understanding a disease's biological mechanisms for a particular individual will allow for improved prognosis and therefore more effective personalized treatments. AI for the purpose of subtyping has already proven to be effective in other fields of medicine, oncology most notably. For example, deep neural networks have consistently demonstrated an ability to classify breast cancer's molecular subtypes [32]. By utilizing biologic data to drive breast cancer's nosology and discovery of disease-modifying treatments, nearly 20 different disease subtypes have been identified, each with a unique response to therapy and survival curve [33].

If PD biomarkers can be developed with a systems-biology model that assumes the disease comprises

patients with several distinct genetic, biological, and pathophysiologic abnormalities, AI can drive innovation in the methods for characterizing them. These powerful AI methods cannot be employed in PD without the requisite molecular data, however. Before a disease's root cause can be identified, researchers must define disease subgroups distinguished by clinical manifestation, posit the mechanisms that could form the subgroups, measure the hypothesized mechanisms, and then use AI to determine the molecular drivers [34]. In this way, disease modification efforts require AI applications not on clinical or biological datasets alone but on combined datasets. Only then will the resulting system outperform what clinicians are able to do today with the clinical data at their disposal.

Efforts to identify biological subgroups in PD, such as those at the Cincinnati Cohort Biomarker Program [35], are nascent but underway. The use of AI-driven clinical and molecular subtyping to guide data collection can enrich molecular-based subtyping efforts and will thus be important to the launch of true precision medicine for PD and other neurodegenerative disorders [36].

CONCLUSION

AI-based tools are embedding into healthcare practice and research. As these algorithms develop, it is necessary to consider the relationship between AI and clinicians for both symptom management and disease modification. The ubiquity of smartphones and inexpensive sensors provides clinicians an unprecedented opportunity to monitor symptoms. Tools designed to capture the patient experience in a free-living environment can enable the personalization of symptomatic treatment to minimize disease burden.

This paper did not intend to be a systematic review of the literature detailing AI's role in diagnosing and managing patients with PD; the choice of examples, papers, and overall conclusions were naturally influenced by the authors' interest in AI and precision medicine. We trust that these biases did not preclude a balanced review of the benefits and challenges associated with leveraging AI. We believe the use of AI for symptom management represents meaningful progress. However, the ultimate frontier for PD remains the discovery of disease-modifying interventions. Because many distinct molecular and biological abnormalities comprise the construct of PD, disease modification efforts demand the

identification of subtypes for which unique therapeutics are suited. The creation of AI algorithms to elucidate the molecular rather than clinical or pathologic nature of PD subtypes would be an important advance in understanding the disease individualized to those affected. This achievement requires an evolution of the clinician's and machine's roles and a shift in the extent to which AI cannot just reproduce but exceed what humans can pursue.

ACKNOWLEDGMENTS

Research reported in this publication was supported by the National Institute of Neurological Disorders and Stroke of the National Institutes of Health under Award Number P50NS108676. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

CONFLICT OF INTEREST

M. Landers has no conflict of interest to report.

S. Saria is a founder of, and holds equity in, Bayesian Health. This arrangement has been reviewed and approved by Johns Hopkins University in accordance with its conflict-of-interest policies. S.S. is a member of the scientific advisory board for PatientPing.

A.J. Espay has no conflict of interest to report.

REFERENCES

- [1] Choi GH, Yun J, Choi J, Lee D, Shim JH, Lee HC, Chung Y-H, Lee YS, Park B, Kim N, Kim KM (2020) Development of machine learning-based clinical decision support system for hepatocellular carcinoma. *Sci Rep* **10**, 14855.
- [2] Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* **316**, 2402.
- [3] Pierson E, Cutler DM, Leskovec J, Mullainathan S, Obermeyer Z (2021) An algorithmic approach to reducing unexplained pain disparities in underserved populations. *Nat Med* **27**, 136-140.
- [4] Lipsmeier F, Taylor KI, Kilchenmann T, Wolf D, Scotland A, Schjodt-Eriksen J, Cheng W, Fernandez-Garcia I, Siebourg-Polster J, Jin L, Soto J, Verselis L, Boess F, Koller M, Grundman M, Monsch AU, Postuma RB, Ghosh A, Kremer T, Czech C, Gossens C, Lindemann M (2018) Evaluation of smartphone-based testing to generate exploratory outcome measures in a phase I Parkinson's disease clinical trial. *Mov Disord* **33**, 1287-1297.

- [5] Postuma RB, Berg D, Stern M, Poewe W, Olanow CW, Oertel W, Obeso J, Marek K, Litvan I, Lang AE, Halliday G, Goetz CG, Gasser T, Dubois B, Chan P, Bloem BR, Adler CH, Deuschl G (2015) MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord* **30**, 1591-1601.
- [6] Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, Poewe W, Sampaio C, Stern MB, Dodel R, Dubois B, Holloway R, Jankovic J, Kulisevsky J, Lang AE, Lees A, Leurgans S, LeWitt PA, Nyenhuis D, Olanow CW, Rascol O, Schrag A, Teresi JA, Hilten JJ van, LaPelle N, Movement Disorder Society UPDRS Revision Task Force (2008) Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Mov Disord* **23**, 2129-2170.
- [7] Dorsey ER, Venuto C, Venkataraman V, Harris DA, Kiebertz K (2015) Novel methods and technologies for 21st-century clinical trials: A review. *JAMA Neurol* **72**, 582-588.
- [8] Coelln R von, Shulman LM (2016) Clinical subtypes and genetic heterogeneity. *Curr Opin Neurol* **29**, 727-734.
- [9] Heim B, Krismer F, Marzi RD, Seppi K (2017) Magnetic resonance imaging for the diagnosis of Parkinson's disease. *J Neural Transm* **124**, 915-964.
- [10] Lim SH, Ferdousi M, Kalteniece A, Mahfoud ZR, Petropoulos IN, Malik RA, Kobylecki C, Silverdale M (2021) Corneal confocal microscopy identifies Parkinson's disease with more rapid motor progression. *Mov Disord*, doi: 10.1002/mds.28602
- [11] Williams S, Fang H, Relton SD, Graham CD, Alty JE (2020) Seeing the unseen: Could Eulerian video magnification aid clinician detection of subclinical Parkinson's tremor? *J Clin Neurosci* **81**, 101-104.
- [12] Zhang J, Cao Y, Li K, Wang F, Wang Y, Wu J, Pei S, Chen J, Mao C, Liu C (2021) Correlations between retinal nerve fiber layer thickness and cognitive progression in Parkinson's disease: A longitudinal study. *Parkinsonism Relat Disord* **82**, 92-97.
- [13] Pablo-Fernández ED, Lees AJ, Holton JL, Warner TT (2019) Clinical Parkinson disease subtyping does not predict pathology. *Nat Rev Neurol* **15**, 189-190.
- [14] Espay AJ, Schwarzschild MA, Tanner CM, Fernandez HH, Simon DK, Leverenz JB, Merola A, Chen-Plotkin A, Brundin P, Kauffman M, Erro R, Kiebertz K, Woo D, Macklin EA, Standaert DG, Lang AE (2017) Biomarker-driven phenotyping in Parkinson's disease: A translational missing link in disease-modifying clinical trials. *Mov Disord* **32**, 319-324.
- [15] Buchman AS, Yu L, Wilson RS, Leurgans SE, Nag S, Shulman JM, Barnes LL, Schneider JA, Bennett DA (2019) Progressive parkinsonism in older adults is related to the burden of mixed brain pathologies. *Neurology* **92**, e1821-e1830.
- [16] Irwin DJ, Grossman M, Weintraub D, Hurtig HI, Duda JE, Xie SX, Lee EB, Deerlin VM, Lopez OL, Kofler JK, Nelson PT, Jicha GA, Woltjer R, Quinn JF, Kaye J, Leverenz JB, Tsuang D, Longfellow K, Yearout D, Kukull W, Keene CD, Montine TJ, Zabetian CP, Trojanowski JQ (2017) Neuropathological and genetic correlates of survival and dementia onset in synucleinopathies: A retrospective analysis. *Lancet Neurol* **16**, 55-65.
- [17] Mei J, Desrosiers C, Frasnelli J (2021) Machine learning for the diagnosis of Parkinson's disease: A systematic review. *Front Aging Neurosci* **13**, 633752.
- [18] Borzì L, Varrecchia M, Olmo G, Artusi CA, Fabbri M, Rizzone MG, Romagnolo A, Zibetti M, Lopiano L (2019) Home monitoring of motor fluctuations in Parkinson's disease patients. *J Reliab Intelligent Environ* **5**, 145-162.
- [19] Arora S, Venkataraman V, Zhan A, Donohue S, Biglan KM, Dorsey ER, Little MA (2015) Detecting and monitoring the symptoms of Parkinson's disease using smartphones: A pilot study. *Parkinsonism Relat Disord* **21**, 650-653.
- [20] Belić M, Bobić V, Badža M, Šolaja N, D*urić-Jovičić M, Kostić VS (2019) Artificial intelligence for assisting diagnostics and assessment of Parkinson's disease—A review. *Clin Neurol Neurosurg* **184**, 105442.
- [21] Tsanas A, Little MA, McSharry PE, Ramig LO (2011) Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity. *J R Soc Interface* **8**, 842-855.
- [22] Perez-Pozuelo I, Zhai B, Palotti J, Mall R, Aupetit M, Garcia-Gomez JM, Taheri S, Guan Y, Fernandez-Luque L (2020) The future of sleep health: A data-driven revolution in sleep science and medicine. *NPJ Digit Med* **3**, 42.
- [23] Zhan A, Mohan S, Tarolli C, Schneider RB, Adams JL, Sharma S, Elson MJ, Spear KL, Glidden AM, Little MA, Terzis A, Dorsey ER, Saria S (2018) Using smartphones and machine learning to quantify Parkinson disease severity: The Mobile Parkinson Disease Score. *JAMA Neurol* **75**, 876.
- [24] Gottlieb S, Statement from FDA Commissioner Scott Gottlieb, M.D., on FDA's efforts to enhance the patient perspective and experience in drug development and review, <https://www.fda.gov/news-events/press-announcements/statement-fda-commissioner-scott-gottlieb-md-fdas-efforts-enhance-patient-perspective-and-experience>, Last updated March 30, 2018, Accessed on December 29, 2020.
- [25] 21st Century Cures Act. H.R. 34, 114th Congress, <https://www.congress.gov/114/plaws/publ255/PLAW-114publ255.pdf>, Last updated December 13, 2016, Accessed on December 29, 2020.
- [26] Riegel B, Moser DK, Glaser D, Carlson B, Deaton C, Armola R, Sethares K, Shively M, Evangelista L, Albert N (2002) The Minnesota Living With Heart Failure Questionnaire. *Nurs Res* **51**, 209-218.
- [27] Healy DG, Falchi M, O'Sullivan SS, Bonifati V, Durr A, Bressman S, Brice A, Aasly J, Zabetian CP, Goldwurm S, Ferreira JJ, Tolosa E, Kay DM, Klein C, Williams DR, Marras C, Lang AE, Wszolek ZK, Berciano J, Schapira AH, Lynch T, Bhatia KP, Gasser T, Lees AJ, Wood NW, International LRRK2 Consortium (2008) Phenotype, genotype, and worldwide genetic penetrance of LRRK2-associated Parkinson's disease: A case-control study. *Lancet Neurol* **7**, 583-590.
- [28] Hood L, Friend SH (2011) Predictive, personalized, preventive, participatory (P4) cancer medicine. *Nat Rev Clin Oncol* **8**, 184-187.
- [29] Brunet JP, Tamayo P, Golub TR, Mesirov JP (2004) Metagenes and molecular pattern discovery using matrix factorization. *Proc Natl Acad Sci U S A* **101**, 4164-4169.
- [30] Perou CM, Sørlie T, Eisen MB, Rijn M van de, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge Ø, Pergamenschikov A, Williams C, Zhu SX, Lønning PE, Børresen-Dale A-L, Brown PO, Botstein D (2000) Molecular portraits of human breast tumours. *Nature* **406**, 747-752.
- [31] Kaissis G, Ziegelmayer S, Lohöfer F, Steiger K, Algül H, Muckenhuber A, Yen H-Y, Rummeny E, Friess H, Schmid R, Weichert W, Siveke JT, Braren R (2019) A machine learning algorithm predicts molecular subtypes in pancreatic ductal adenocarcinoma with differential response

- to gemcitabine-based versus FOLFIRINOX chemotherapy. *PLoS One* **14**, e0218642.
- [32] Mohaiminul Islam M, Huang S, Ajwad R, Chi C, Wang Y, Hu P (2020) An integrative deep learning framework for classifying molecular subtypes of breast cancer. *Comput Struct Biotechnol J* **18**, 2185-2199.
- [33] Malhotra GK, Zhao X, Band H, Band V (2010) Histological, molecular and functional subtypes of breast cancers. *Cancer Biol Ther* **10**, 955-960.
- [34] Saria S, Goldenberg A (2015) Subtyping: What it is and its role in precision medicine. *IEEE Intell Syst* **30**, 70-75.
- [35] Sturchio A, Marsili L, Vizcarra JA, Dwivedi AK, Kauffman MA, Duker AP, Lu P, Pauciulo MW, Wissel BD, Hill EJ, Stecher B, Keeling EG, Vagal AS, Wang L, Haslam DB, Robson MJ, Tanner CM, Hagey DW, Andaloussi SE, Ezzat K, Fleming RMT, Lu LJ, Little MA, Espay AJ (2020) Phenotype-agnostic molecular subtyping of neurodegenerative disorders: The Cincinnati Cohort Biomarker Program (CCBP). *Front Aging Neurosci* **12**, 553635.
- [36] Dorsey ER, Omberg L, Waddell E, Adams JL, Adams R, Ali MR, Amodeo K, Arky A, Augustine EF, Dinesh K, Hoque ME, Glidden AM, Jensen-Roberts S, Kabelac Z, Katabi D, Kiebertz K, Kinel DR, Little MA, Lizarraga KJ, Myers T, Riggare S, Rosero SZ, Saria S, Schifitto G, Schneider RB, Sharma G, Shoulson I, Stevenson EA, Tarolli CG, Luo J, McDermott MP (2020) Deep phenotyping of Parkinson's disease. *J Parkinsons Dis* **10**, 855-873.