Received: 27 November 2019 | Revised: 21 February 2020 | Accepted: 6 April 2020

DOI: 10.1002/pds.5015

WILEY

ORIGINAL REPORT

Identifying drugs with disease-modifying potential in Parkinson's disease using artificial intelligence and pharmacoepidemiology

Laura C. Maclagan[1] | Naomi P. Visanji[2] | Yi Cheng[1] | Mina Tadrous[1,3] | Alix M. B. Lacoste[4] | Lorraine V. Kalia[2] | Susan E. Bronskill[1,3,5] | Connie Marras[1,2]

[1]ICES, Life Stage Research Program, Toronto, Ontario, Canada

[2]Edmond J Safra Program in Parkinson Disease, Toronto Western Hospital, Toronto, Ontario, Canada

[3]Women's College Research Institute, Toronto, Ontario, Canada

[4]Data Science, BenevolentAI, Brooklyn, New York

[5]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada

Correspondence
Connie Marras, Edmond J Safra Program in Parkinson Disease, Toronto Western Hospital, 399 Bathurst St, Toronto, ON M5T 2S8, Canada.
Email: connie.marras@uhnresearch.ca

Funding information
Institute for Clinical Evaluative Sciences (ICES); Ontario Brain Institute; Ontario Ministry of Health and Long-TermCare; Ontario Neurodegenerative Disease Research Initiative (ONDRI)

Abstract

Purpose: The aim of the study was to assess the feasibility of an approach combining computational methods and pharmacoepidemiology to identify potentially disease-modifying drugs in Parkinson's disease (PD).

Methods: We used a two-step approach; (a) computational method using artificial intelligence to rank 620 drugs in the Ontario Drug Benefit formulary based on their predicted ability to inhibit alpha-synucleinaggregation, a pathogenic hallmark of PD; and (b) case-control study using administrative databases in Ontario, Canada. Persons aged 70-110 years with incident PD from April 2002-March 2013. Controls were randomly selected from persons with no previous diagnosis of PD.

Results: A total of 15 of the top 50 drugs were deemed feasible for pharmacoepidemiologic analysis, of which seven were significantly associated with incident PD after adjustment, with five of these seven associated with a decreased odds of PD. Methylxanthine drugs pentoxifylline (OR, 0.72; 95% CI, 0.59-0.89) and theophylline (OR, 0.77; 95% CI, 0.66-0.91), and the corticosteroid dexamethasone (OR, 0.72; 95% CI, 0.61-0.85) were associated with decreased odds of PD.

Conclusions: Our findings demonstrate the feasibility of this approach to focus the search for disease-modifying drugs. Corticosteroids and methylxanthines should be further investigated as potential disease-modifyingdrugs in PD.

KEYWORDS

artificial intelligence, drug repositioning, Parkinson disease, pharmacoepidemiology

1 | INTRODUCTION

Parkinson's disease (PD) is an age-related neurodegenerative disease affecting more than 10 million individuals worldwide, a number anticipated to double between 2005 and 2030.[1,2] Presently, treatment of PD is focused on dopamine replacement therapy, with more invasive surgical interventions available for management of advanced cases.[3] These therapies, while effective at managing motor symptoms, do not prevent or slow the pathogenic processes underlying the disease's continuing neurodegeneration. The lack of disease-modifyingtreatment able to delay progression of the neurodegenerative process in PD remains a critical unmet need.[4]

The cause of neurodegeneration in PD is multifaceted;[5] however, a defining feature is aggregation of the protein alpha synuclein (aSyn) in characteristic inclusions in brain regions exhibiting neurodegeneration.[6,7] A growing number of preclinical studies have

Maclagan and Visanji are co-first authors.

Bronskill and Marras are co-senior authors.

Maclagan, Visanji, Bronskill, and Marras contributed equally to the work.

demonstrated that aggregated forms of aSyn are inherently neurotoxic and are capable of spreading throughout the brain and inducing neurodegeneration.[8-10] Therefore, preventing aSyn aggregation has become the focus of many drug development efforts seeking a potential disease-modifying therapy for PD.[11]

Drug development is a time-consuming, expensive and inefficient process[12-14] and, with a growing list of failed clinical trials, pharmaceutical companies are beginning to abandon research efforts to develop a novel disease-modifying treatment for PD.[15] By comparison, drug repurposing, taking an already approved drug and developing it for a new indication, can be an efficient method of providing new treatments for disease.[13] A repurposed drug can enter clinical trials for a new indication at Phase IIa, making the development process significantly less time-consuming, less expensive and with a much higher probability of success.[13,16] The probability of success for a repurposed drug was recently estimated at ~33%.[17] Although clearly an attractive method of developing new treatments for unmet medical needs, a major challenge in drug repurposing is identifying a drug with efficacy for the indication of interest from the many thousands of approved drugs.[18,19]

To address this challenge, we have harnessed the power of a computational approach including artificial intelligence to make predictions about the potential efficacy of existing drugs to reduce the aggregation of aSyn and thus have disease-modifying potential in PD. Using this same methodology, we have previously identified drugs that could be repurposed to treat levodopa-induced dyskinesia in PD.[18] We assessed the feasibility of combining this computational method (to rank candidate drugs that may be disease-modifying in PD), and pharmacoepidemiology (to validate these hypotheses using real-worldhealth administrative databases) in a two-step approach to identify potentially disease-modifying drugs for PD.

KEY POINTS

- To our knowledge, this is the first study to assess the feasibility of a combined artificial intelligence and pharmacoepidemiology approach to identify potentially disease-modifying drugs.
- We ranked a list of 620 drugs using a computational approach based on predicted ability to inhibit alpha synuclein aggregation, a pathological process associated with PD.
- A total of 15 of the top 50 drugs screened were suitable for pharmacoepidemiologic analyses using a case-control design.
- Seven drugs were significantly associated with incident PD. We identified methylaxthine drugs pentoxifylline and theophylline as novel disease-modifying agents.
- This approach may be used to rapidly to identify drugs suitable for repurposing.

2 | METHODS

2.1 | Step 1: Computational approach to identify drugs with predicted disease-modifying potential in PD

We used IBM Watson to rank all 620 unique active compounds in the Ontario Drug Benefit (ODB) formulary based on semantic similarity to a set of 15 compounds that have been demonstrated to reduce aggregation of aSyn in cell-basedor animal models (Table S1 in Data S1).[20] The approach to ranking first comprises a text mining step to generate a mathematical representation (vector) of each drug based on semantic information extracted from all PubMed abstracts mentioning each drug up to the date of analysis (20 October 2016). These vectors represent words and phrases that co-occurwith the corresponding drugs in text. They are used to generate a distance matrix that contains a similarity index for every individual drug included in the analysis when paired with every other individual drug (vector space model). Finally, a network diffusion algorithm (graph diffusion) was applied to rank order each candidate based on similarity to the set of 15 known compounds. This ranking resulted in a list of candidate drugs ordered by predicted ability to reduce aggregation of aSyn and thus have disease-modifying potential in PD. A full description of the ranking model has been provided previously.[21]

2.2 | Step 2: Pharmacoepidemiology using health administrative databases

To examine associations between potentially disease-modifying drug exposures (identified in step 1) and PD in a real-world setting, we performed a case-control validation study using health administrative databases covering the population of Ontario, Canada. Ontario has a population of approximately 14 million people covered under universal health insurance. Coverage is provided for medically necessary physician and hospital services for all ages, and for medications for persons aged 65 and older and those receiving social assistance. A complete listing of health administrative databases utilized in this study can be found in Table S2 in Data S1. These datasets were linked using unique encoded identifiers and analyzed at ICES.

2.3 | Defining PD cases and controls

Persons with incident PD were identified using a validated algorithm of three physician visits for PD within a 2-year period (with a minimum of a 30-day gap between each), using methods previously described.[22] The definition was validated using clinician review of electronic medical records as the reference standard and was found to have a sensitivity of 72.3% (95% CI, 65.9%-78.6%) and a specificity

of 100.0% (95% CI, 99.9%-100.0%) (Butt et al, unpublished work). The date of PD case ascertainment (index date) was the date of the first physician visit meeting the case definition. Persons meeting the PD case definition were accrued from April 1, 2002 to March 31, 2013 and were required to be aged 70-110 years on study index. Cases were matched to 5 controls who had no previous diagnosis of PD (as of the index date) using incidence density sampling from among persons alive and eligible for health insurance on the PD case ascertainment date (thus controlling for the effect of calendar time). Cases and controls were matched on age (±1 year), sex, and comorbidity level as measured using Aggregated Diagnosis Groups (±1 ADGs). ADGs were derived using the Johns Hopkins Adjusted Clinical Group (ACG) System Version 10.0 (https://www.hopkinsacg.org/) using diagnoses in hospitalization records and physician claims 2 years prior to index. Cases and controls were excluded, if they were not eligible for Ontario health insurance on their index date, or if they had a previous diagnosis of dementia which could confound diagnosis of PD.

2.4 | Candidate drug screening

To qualify as feasible for the pharmacoepidemiologic analysis, a drug must have been a) covered by the formulary and frequently dispensed (>100 claims per quarter) during the drug exposure period (5 years prior to study index date) and b) mostly acquired by prescription rather than over the counter. Non-steroidalanti-inflammatorymedications (NSAIDs) were excluded due to known associations with PD and substantial over the counter use.[23] Topically administered medications were excluded based on questionable biological plausibility of disease-modifyingeffect for PD. Drugs available as combination products only were also excluded. Medications for dementia were excluded given their use to treat cognitive symptoms of PD. One criterion was applied to the PD cohort: sufficient frequency of use in the PD cohort (>0.2%).

2.5 | Characteristics of cases and controls at baseline

Baseline sociodemographic characteristics including age, sex, neighborhood income quintile, urban/rural residence, nursing home residence, and low income were assessed at the study index date. Prevalent chronic conditions at index were identified using validated health administrative data algorithms that have been extensively used in previous studies.[24-29] Health care utilization in the year prior to index was assessed, including hospitalizations, visits to general practitioners/family physicians, neurologists, geriatricians, and other specialists (those not included in previous groups). The number of unique medications (by drug name) was also assessed in the year prior to index. Anti-parkinsonmedications included levodopa containing agents (Table S3 in Data S1).

2.6 | Statistical analyses

Standardized differences were used to compare baseline characteristics between persons with PD and controls, as they are not influenced by sample size.[30] Standardized differences <0.10 were considered well-balanced. Duration(s) of use was calculated using days supplied over the entire drug exposure period. Gaps in days

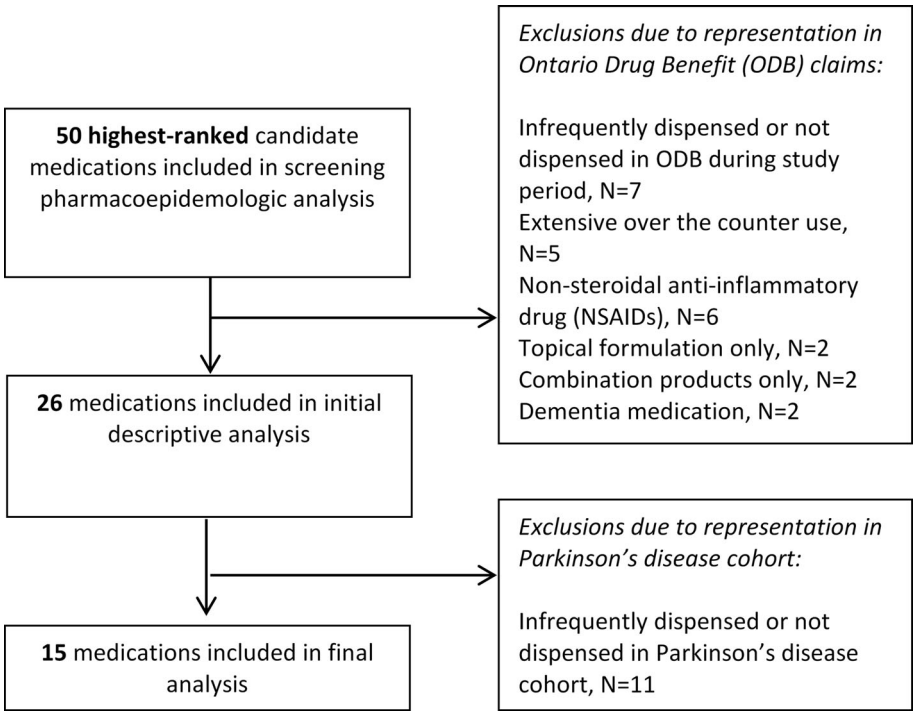

FIGURE 1 Flow diagram for drug selection

TABLE 1 Baseline characteristics of Parkinson's disease cases and controls, Ontario, Canada

Characteristics, n (%)[a]	Cases N = 14 866	Controls N = 74 330	Standardized difference
Demographics			
Age, y			
Mean (SD)	78.2 (5.4)	78.0 (5.4)	0.02
Age groups			
70-74	4341 (29.2%)	22 771 (30.6%)	0.03
75-79	4844 (32.6%)	23 733 (31.9%)	0.01
80-84	3667 (24.7%)	17 889 (24.1%)	0.01
85+	2014 (13.5%)	9937 (13.4%)	0.01
Male sex	8381 (56.4%)	41 905 (56.4%)	0
Low income senior	3026 (20.4%)	16 864 (22.7%)	0.06
Neighborhood income quintile	60 (0.4%)	205 (0.3%)	0.02
1 (lowest)	2619 (17.6%)	14 163 (19.1%)	0.04
2	2966 (20.0%)	15 756 (21.2%)	0.03
3	2903 (19.5%)	14 608 (19.7%)	0
4	3027 (20.4%)	14 574 (19.6%)	0.02
5 (highest)	3291 (22.1%)	15 024 (20.2%)	0.05
Rural residence	1942 (13.1%)	10 274 (13.8%)	0.02
Nursing home resident	350 (2.4%)	956 (1.3%)	0.08
Health status			
Diabetes	3871 (26.0%)	20 236 (27.2%)	0.03
Hypertension	10 835 (72.9%)	55 917 (75.2%)	0.05
Asthma	1756 (11.8%)	9473 (12.7%)	0.03
COPD	**1353 (9.1%)**	**9258 (12.5%)**	**0.11**
CHF	1685 (11.3%)	9891 (13.3%)	0.06
Rheumatoid arthritis	295 (2.0%)	1739 (2.3%)	0.02
AMI	1141 (7.7%)	7108 (9.6%)	0.07
Stroke/Transient ischemic attack	815 (5.5%)	4145 (5.6%)	0
Peripheral vascular disease	396 (2.7%)	2942 (4.0%)	0.07
Chronic kidney disease	1170 (7.9%)	6970 (9.4%)	0.05
Aggregated diagnosis groups (ADGs)			
0-5	2940 (19.8%)	15 323 (20.6%)	0.02
6-10	7215 (48.5%)	36 396 (49.0%)	0.01
10+	4711 (31.7%)	22 611 (30.4%)	0.03
Medication use (in the year prior to index)			
Unique medications (mean ± SD)	**9.37 ± 5.49**	**8.72 ± 5.51**	**0.12**
Receipt of anti-Parkinsonmedication	**6472 (43.5%)**	**524 (0.7%)**	**1.2**
Health care utilization (in year prior to index)			
At least one hospitalization	2825 (19.0%)	14 161 (19.1%)	0
At least one FP/GP visit	14 289 (96.1%)	71 197 (95.8%)	0.02
At least one neurologist visit	**2652 (17.8%)**	**3191 (4.3%)**	**0.44**
At least one geriatrician visit	**421 (2.8%)**	**650 (0.9%)**	**0.15**
At least one psychiatrist visit	**419 (2.8%)**	**730 (1.0%)**	**0.13**
At least one other specialist visit	11 850 (79.7%)	59 033 (79.4%)	0.01

[a]Unless otherwise noted. Standardized differences >0.10 indicating imbalance are bolded.

supplied were permitted. Where medication dispensing records overlapped, days supplied were counted once. Conditional logistic regression models were used to assess associations between previous exposure to a potentially disease-modifyingdrug and incident PD. Unadjusted models and adjusted models were constructed including adjustment for potential confounders suspected to be associated with PD and with indications for potentially disease-modifyingdrugs. Confounders included urban/rural residence, income quintile, low-incomesenior, long-termcare resident, asthma, hypertension, diabetes, chronic obstructive pulmonary disease, previous acute myocardial infarction, stroke/transient ischemic attack, congestive heart failure, renal disease, peripheral vascular disease and number of medications in the past year. All analyses were performed using SAS Version 9.4 (SAS Institute, Cary, NC, USA).

2.7 | Sensitivity analyses

In a sensitivity analysis, candidate drug exposure in the 4 years immediately prior to index was excluded to minimize the likelihood that prescribing was addressing symptoms during the prodromal disease period prior to diagnosis (ie, protopathic bias). The association between incident PD and duration of exposure to candidate drugs was assessed using models categorized in tertiles of days supplied for each candidate drug (0 days supplied [reference], >0 to 33rd percentile, >33rd to 66th percentile ad 66th percentile to maximum) (Table S4 in Data S1).

3 | RESULTS

The top 50 of these 620 ranked candidate drugs are presented in Table S5 in Data S1. After exclusions based on ODB claims criteria were applied, 26 candidate drugs remained out of the top 50. In pharmacoepidemiologic analysis, 14 866 persons with incident PD (cases) were identified during the study period. The candidate drug list was further reduced to 15 candidate drugs for pharmacoepidemiologic evaluation, when PD cohort exclusion criteria were applied (Figure 1).

3.1 | Characteristics of PD cases and controls

Each case was matched to five eligible controls yielding 74 330 controls with no previous diagnosis of PD. The cases and controls were generally well-balancedin their sociodemographic, health status, health care and medication use characteristics (Table 1). Cases and controls were well-balancedon age and drug exposure period prior to index (mean: 13.2 years in cases vs 13.0 years in controls, standardized difference = 0.02).

Cases were less likely to have chronic obstructive pulmonary disease as compared to controls, consistent with the known inverse association between PD and smoking. Cases received higher numbers of medications and were more likely to have received an anti-parkinsonmedication in the year prior to index. Cases were more likely to have had a neurologist, psychiatrist and geriatrician visit in the year prior to index as compared to controls but were equally likely to visit other specialists.

3.2 | Frequency of candidate disease-modifyingdrug use

The most commonly prescribed candidate disease-modifyingdrug was estradiol (15% of cases and 13.7% of controls, Table 2). Least commonly prescribed medications were hydrocortisone (0.4% of cases and controls) and ketoconazole (0.8% of cases and 0.7% of controls).

3.3 | Validation analysis

In both unadjusted logistic regression models and models adjusted for relevant confounders, exposure to allopurinol, dexamethasone, fenofibrate, pentoxifylline and theophylline was associated with a decreased odds of incident PD (Figure 2, see Table S6 in Data S1 for unadjusted and adjusted estimates). Previous exposure to estradiol and propranolol was associated with a significantly increased odds of PD. Other drugs examined did not show significant associations with incident PD.

TABLE 2 Previous exposure to potentially disease-modifyingdrugs in persons with incident PD and controls

Potentially disease-modifying drug	Cases N = 14 866	Controls N = 74 330	Standardized difference
Allopurinol	1013 (6.8%)	6085 (8.2%)	0.05
Dexamethasone	155 (1.0%)	982 (1.3%)	0.03
Estradiol	2225 (15.0%)	10 158 (13.7%)	0.04
Fenofibrate	634 (4.3%)	3424 (4.6%)	0.02
Hydrocortisone	54 (0.4%)	298 (0.4%)	0.01
Ketoconazole	114 (0.8%)	502 (0.7%)	0.01
Losartan	717 (4.8%)	3650 (4.9%)	0
Lovastatin	463 (3.1%)	2455 (3.3%)	0.01
Pentoxifylline	111 (0.7%)	768 (1.0%)	0.03
Pioglitazone	176 (1.2%)	909 (1.2%)	0
Propranolol	**1143 (7.7%)**	**1611 (2.2%)**	**0.26**
Tamoxifen	172 (1.2%)	829 (1.1%)	0
Tetracycline	975 (6.6%)	4447 (6.0%)	0.02
Theophylline	180 (1.2%)	1192 (1.6%)	0.03
Verapamil	426 (2.9%)	2247 (3.0%)	0.01

Standardized differences >0.10 indicating imbalance are bolded.

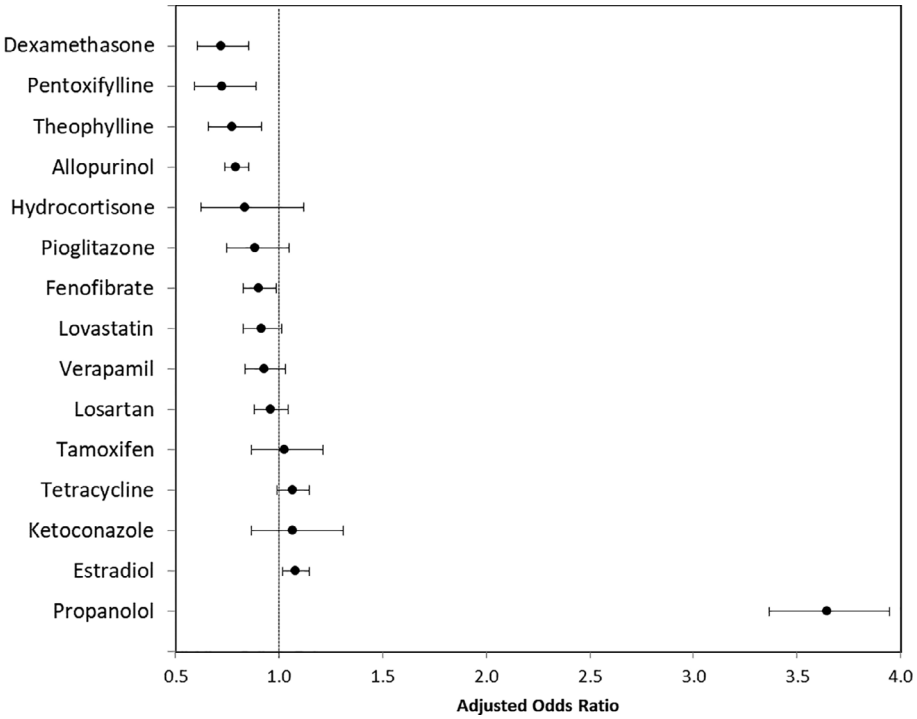

FIGURE 2 Odds of incident Parkinson's disease associated with previous exposure to potentially disease-modifyingmedications. Drugs are listed in ascending order of odds ratio effect estimate. Error bars indicate 95% confidence intervals surrounding odds ratio point estimates. Adjusted for urban/rural residence, income quintile, low-incomesenior, nursing home resident, asthma, hypertension, diabetes, COPD, previous MI, stroke/TIA, CHF, renal disease, peripheral vascular disease, number of medications in past year

3.4 | Sensitivity analyses

In sensitivity analyses, protective associations with allopurinol and theophylline were again observed, when the four-yeardrug exposure period immediately prior to index date was excluded (Table S7 in Data S1). An increased odds of PD associated with estradiol and propranolol were also observed, albeit with a smaller effect estimate for propranololthan in the primary analysis. Associations with dexamethasone, fenofibrate and pentoxifylline were not significant; however, pentoxifylline showed a non-significanttrend toward a protective effect.

Longer durations of dexamethasone exposure were associated with lower odds of incident PD (Table S8 in Data S1). The shortest tertiles of lovastatin and theophylline use were also associated with a significantly reduced odds of PD. Longer durations of pentoxifylline exposure were associated with reduced odds of PD, and effect estimates showed a trend toward dose-response. Propanolol was significantly associated with PD at all levels of exposure but showed the largest effect estimates for the shortest and medium durations.

4 | DISCUSSION

In the present study, we identified existing drugs that may be further studied for repurposing as disease-modifyingtherapies in PD. Our pairing of a computational approach to rank potentially disease-modifyingdrugs based on inhibition of aSyn aggregation with a real-worldpharmacoepidemiology validation in a large cohort of almost 15 000 persons with incident PD is a feasible and novel method for assessing the role of repurposed drugs in PD in a real-worldcontext. We identified potentially disease modifying drugs in PD that may be prioritized for future research, some of which were expected based on previously published findings (highlighting the reliability of our data), whereas others represent novel and potentially exciting findings.

We found an expected protective association with the urate-loweringmedication allopurinol, consistent with the known association of PD with lower levels of serum uric acid.[31] Cholesterol-loweringdrugs fenofibrate and lovastatin were also inversely associated with PD, consistent with some prior studies showing a protective effect of statins on PD, but may be a reflection of the inverse association between serum cholesterol and PD.[32] Also expected was a positive association between propranolol and PD, likely a reflection of use of propranolol to control tremor as an early symptom of PD prior to diagnosis. Although propranolol is not a common treatment for classical resting tremor in PD, it may be prescribed when the tremor is not recognized as a symptom of PD (a common situation when tremor appears first as the only symptom). The fact that only shorter durations of propranolol exposure were significantly associated with PD is supportive of this hypothesis. It has also been suggested that beta blockers may exacerbate the degenerative process in PD related to increased alpha synuclein gene expression.[33] Our findings are consistent with this hypothesis, although our study cannot determine the cause of the association. Finding these expected associations suggests that our measurement of drug exposure and outcome (PD) perform well and enable us to detect associations when present.

The inverse association with the most convincing dose-responserelationship was between PD and the corticosteroid dexamethasone. This is consistent with a recent large population-based-case-controlstudy finding that corticosteroids were associated with a

lower risk of PD.[34] Having also found another immunosuppressant medication class, inosine monophosphate dehydrogenase inhibitors, also associated with reduced PD risk the authors proposed that immune suppressing effects specifically on T-cellscould mediate protective effects against PD.

Unexpected and, to our knowledge, novel inverse associations were identified between PD and both pentoxifylline and theophylline. A suggestive dose-responserelationship was found for pentoxifylline. Pentoxifylline and theophylline both belong to the methylxanthine class of drugs, which are known to have immunomodulatory effects.[35] Methylxanthines are phosphodiesterase inhibitors which inhibit synthesis of pro-inflammatorymediators such as tumor necrosis factor and leukotrienes.[36-39] Collectively these effects are anti-inflammatoryand, as neuroinflammation plays a critical role in the pathogenesis of PD,[40] an anti-inflammatorymechanism provides a tempting explanation for our findings. Other anti-inflammatorydrugs, such as ibuprofen, have also been associated with a reduced risk of PD.[23] It is notable that caffeine, well-knownto be inversely associated with PD risk, is also a methylxanthine derivative. Methylxanthines are also nonselective antagonists at adenosine receptors. It has been suggested that adenosine A2a inhibitors may have anti-inflammatoryand neuroprotective effects in PD[41,42] and may also have symptomatic benefit in PD.[43] Theophylline has been suggested to have some symptomatic benefit in advanced PD patients, presumed to be through its ability to potentiate the effects of levodopa.[44] Importantly, no significant symptomatic effect was seen in early levodopa-naïve PD patients, which would be the population most relevant to our findings. Thus, it seems unlikely that symptom benefit in the earliest stages of PD accounts for our findings.

It is important to consider the role of smoking in these results. Corticosteroids, pentoxifylline and theophylline are all drugs that are used in smoking-relateddiseases, which is relevant considering the known inverse association between smoking and incident PD. Smoking data are not available in the Ontario administrative databases; however, our analyses adjusted for prior diagnosis of COPD, which is correlated with smoking and highly related to corticosteroid and theophylline use. Our finding of a positive association with estradiol exposure was unexpected; postmenopausal hormone replacement therapy and estrogen specifically have been associated with null or inverse associations with PD in prior studies.[45,46] However, this association was only present for short duration exposure, which suggests that it is unlikely to be based on a biological effect. A possible explanation may be increased health service contact around the time of PD diagnosis resulting in the initiation of other general health maintenance measures.

Strengths of our study include the evaluation of the top 50 ranked candidate drugs using computational methods, a complementary pharmacoepidemiologic validation using large health administrative databases allowing us to detect small but potentially biologically important associations, and the population-basednature of the work enhancing the generalizability of the results to the older adult population. Limitations of our work include the lack of confirmation of PD diagnosis in the administrative data and a reliance on diagnostic coding for reimbursement purposes. Nevertheless, the fact that we were able to reproduce known associations supports the validity of our approach. Our analysis was also restricted to individuals aged 70 +, since drug dispensing information is only available for persons aged 65+. Therefore, the observed associations with PD warrant further study in younger individuals. Related to this, we were unable to document early or early to mid-lifeexposure to the drugs of interest, which may have important effects on the disease process. Although the present study examined associations with the top 50 candidate drugs in order to focus on epidemiologic investigations, future work could build on this approach to examine all medications using joint artificial intelligence and pharmacoepidemiologic approaches.

5 | CONCLUSION

We have demonstrated the efficiency and feasibility of a novel two-stepcomputational and pharmacoepidemiologic approach to identify priority candidates for further study as disease-modifyingagents in PD. Of note, this approach can easily be extended to other mechanisms related to PD pathology besides aSyn aggregation and to other diseases. Given the inverse associations with incident PD observed in the present study, corticosteroids and methylxanthines have emerged as novel potentially disease-modifyingdrugs in PD worthy of further study. Logical next steps would be confirming the pentoxifylline and theophylline inverse associations in other large databases and in other populations, examining the association of these drug classes as well as corticosteroids with disease progression milestones in PD and investigating the effects of these drugs in cellular and animal models of aSyn aggregation.

ETHICS STATEMENT

The use of data in this project was authorized under section 45 of Ontario's Personal Health Information Protection Act, which does not require review by a Research Ethics Board.

ACKNOWLEDGEMENTS

This research was conducted with support from the Ontario Neurodegenerative Disease Research Initiative (ONDRI) through the Ontario Brain Institute, an independent non-profitcorporation, funded partially by the Ontario government. This study was supported by ICES, which is funded by an annual grant from the Ontario Ministry of Health and Long-TermCare (MOHLTC). Parts of this material are based on data and/or information compiled and provided by CIHI. The opinions, results and conclusions are those of the authors, and no endorsement by the Ontario Brain Institute, the MOHLTC, or CIHI is intended or should be inferred. The funders had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review or approval of the manuscript; and decision to submit the manuscript for publication. We thank IMS Brogan Inc. for use of their Drug Information Database.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

The dataset from this study is held securely in coded form at ICES. While data sharing agreements prohibit ICES from making the dataset publicly available, access may be granted to those who meet pre-specifiedcriteria for confidential access, available at www.ices.on.ca/DAS. The full dataset creation plan and underlying analytic code are available from the authors upon request, understanding that the computer programs may rely upon coding templates or macros that are unique to ICES and are therefore either inaccessible or may require modification.

ORCID

Laura C. Maclagan https://orcid.org/0000-0002-2355-6939
Naomi P. Visanji https://orcid.org/0000-0001-5968-7845

REFERENCES

1. Dorsey ER, Constantinescu R, Thompson JP, et al. Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. *Neurology*. 2007;68(5):384-386.
2. Pringsheim T, Jette N, Frolkis A, Steeves TD. The prevalence of Parkinson's disease: a systematic review and meta-analysis. *Mov Disord*. 2014;29(13):1583-1590.
3. Poewe W, Seppi K, Tanner CM, et al. Parkinson disease. *Nat Rev Dis Primers*. 2017;3:17013.
4. Macleod AD, Taylor KS, Counsell CE. Mortality in Parkinson's disease: a systematic review and meta-analysis. *Mov Disord*. 2014;29(13):1615-1622.
5. Johnson ME, Stecher B, Labrie V, Brundin L, Brundin P. Triggers, facilitators, and aggravators: redefining Parkinson's disease pathogenesis. *Trends Neurosci*. 2019;42(1):4-13.
6. Goedert M, Spillantini MG, Del Tredici K, Braak H. 100 years of Lewy pathology. *Nat Rev Neurol*. 2013;9(1):13-24.
7. Spillantini MG, Schmidt ML, Lee VM, Trojanowski JQ, Jakes R, Goedert M. α-synucleinin Lewy bodies. *Nature*. 1997;388(6645):839-840.
8. Braak H, Del Tredici K, Rub U, de Vos RA, Jansen Steur EN, Braak E. Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol Aging*. 2003;24(2):197-211.
9. Ingelsson M. Alpha-synucleinoligomers-neurotoxicmolecules in Parkinson's disease and other Lewy body disorders. *Front Neurosci*. 2016;10:408.
10. Karpowicz RJ Jr, Trojanowski JQ, Lee VM. Transmission of α-synucleinseeds in neurodegenerative disease: recent developments. *Lab Invest*. 2019;99(7):971-981.
11. Stefanis L. α-synucleinin Parkinson's disease. *Cold Spring Harb Perspect Med*. 2012;2(2):a009399.
12. Scannell JW, Blanckley A, Boldon H, Warrington B. Diagnosing the decline in pharmaceutical R&D efficiency. *Nat Rev Drug Discov*. 2012;11(3):191-200.
13. Nosengo N. Can you teach old drugs new tricks? *Nature*. 2016;534:314-316.
14. Pammolli F, Magazzini L, Riccaboni M. The productivity crisis in pharmaceutical R&D. *Nat Rev Drug Discov*. 2011;10(6):428-438.
15. Hawkes N. Pfizer abandons research into Alzheimer's and Parkinson's diseases. *BMJ*. 2018;360:k122.
16. Beachy SH, Johnson SG, Olson SB, Berger AC. *Drug Repurposing and Repositioning. Roundtable on Translating Genomic-BasedResearch for Health. Workshop Summary*. Washington, DC: The National Academies Press; 2014.
17. Neuberger A, Oraiopoulos N, Drakeman DL. Renovation as innovation: is repurposing the future of drug discovery research? *Drug Discov Today*. 2019;24(1):1-3.
18. Johnston TH, Lacoste AMB, Visanji NP, Lang AE, Fox SH, Brotchie JM. Repurposing drugs to treat l-DOPA-induceddyskinesia in Parkinson's disease. *Neuropharmacology*. 2019;147:11-27.
19. Pushpakom S, Iorio F, Eyers PA, et al. Drug repurposing: progress, challenges and recommendations. *Nat Rev Drug Discov*. 2019;18(1):41-58.
20. Visanji N, Lacoste A, Spangler S, et al. In silico predictive analytics: accelerating identification of potentialdisease-modifying compounds for Parkinson's disease [abstract]. *Mov Disord*. 2017;32(suppl 2).
21. Bakkar N, Kovalik T, Lorenzini I, et al. Artificial intelligence in neurodegenerative disease research: use of IBM Watson to identify additional RNA-bindingproteins altered in amyotrophic lateral sclerosis. *Acta Neuropathol*. 2018;135(2):227-247.
22. Butt DA, Tu K, Young J, et al. A validation study of administrative data algorithms to identify patients with Parkinsonism with prevalence and incidence trends. *Neuroepidemiology*. 2014;43(1):28-37.
23. Gagne JJ, Power MC. Anti-inflammatorydrugs and risk of Parkinson disease: a meta-analysis. *Neurology*. 2010;74(12):995-1002.
24. Schultz SE, Rothwell DM, Chen Z, Tu K. Identifying cases of congestive heart failure from administrative data: a validation study using primary care patient records. *Chronic Dis Inj Can*. 2013;33(3):160-166.
25. Lipscombe LL, Hwee J, Webster L, Shah BR, Booth GL, Tu K. Identifying diabetes cases from administrative data: a population-basedvalidation study. *BMC Health Serv Res*. 2018;18(1):316.
26. Tu K, Campbell NR, Chen ZL, Cauch-Dudek KJ, McAlister FA. Accuracy of administrative databases in identifying patients with hypertension. *Open Med*. 2007;1(1):e18-e26.
27. Tu JV, Chu A, Donovan LR, et al. The Cardiovascular Health in Ambulatory Care Research Team (CANHEART): using big data to measure and improve cardiovascular health and healthcare services. *Circ Cardiovasc Qual Outcomes*. 2015;8(2):204-212.
28. Widdifield J, Bombardier C, Bernatsky S, et al. An administrative data validation study of the accuracy of algorithms for identifying rheumatoid arthritis: the influence of the reference standard on algorithm performance. *BMC Musculoskelet Disord*. 2014;15:216.
29. Gershon AS, Wang C, Guan J, Vasilevska-Ristovska J, Cicutto L, To T. Identifying individuals with physcian diagnosed COPD in health administrative databases. *COPD*. 2009;6(5):388-394.
30. Austin PC. The relative ability of different propensity score methods to balance measured covariates between treated and untreated subjects in observational studies. *Medical Decis Mak*. 2009;29(6):661-677.
31. Wen M, Zhou B, Chen YH, et al. Serum uric acid levels in patients with Parkinson's disease: a meta-analysis. *PloS One*. 2017;12(3):e0173731.
32. Huang X, Abbott RD, Petrovitch H, Mailman RB, Ross GW. Low LDL cholesterol and increased risk of Parkinson's disease: prospective results from Honolulu-AsiaAging Study. *Mov Disord*. 2008;23(7):1013-1018.
33. Mittal S, Bjornevik K, Im DS, et al. β2-Adrenoreceptoris a regulator of the α-synucleingene driving risk of Parkinson's disease. *Science*. 2017;357(6354):891-898.
34. Racette BA, Gross A, Vouri SM, Camacho-Soto A, Willis AW, Searles NS. Immunosuppressants and risk of Parkinson disease. *Ann Clin Transl Neurol*. 2018;5(7):870-875.
35. Entzian P, Bitter-Suermann S, Burdon D, Ernst M, Schlaak M, Zabel P. Differences in the anti-inflammatoryeffects of theophylline and pentoxifylline: important for the development of asthma therapy? *Allergy*. 1998;53(8):749-754.
36. Peters-Golden M, Canetti C, Mancuso P, Coffey MJ. Leukotrienes: underappreciated mediators of innate immune responses. *J Immunol*. 2005;174(2):589-594.
37. Deree J, Martins JO, Melbostad H, Loomis WH, Coimbra R. Insights into the regulation of TNF-α production in human mononuclear cells:

the effects of non-specificphosphodiesterase inhibition. *Clinics*. 2008; 63(3):321-328.
38. Marques LJ, Zheng L, Poulakis N, Guzman J, Costabel U. Pentoxifylline inhibits TNF- α production from human alveolar macrophages. *Am J Respir Crit Care Med*. 1999;159(2):508-511.
39. Howell RE. Multiple mechanisms of xanthine actions on airway reactivity. *J Pharmacol Exp Ther*. 1990;255(3):1008-1014.
40. De Virgilio A, Greco A, Fabbrini G, et al. Parkinson's disease: autoimmunity and neuroinflammation. *Autoimmun Rev*. 2016;15(10):1005-1011.
41. Ohta A, Sitkovsky M. Role of G-protein-coupledadenosine receptors in downregulation of inflammation and protection from tissue damage. *Nature*. 2001;414(6866):916-920.
42. Ross GW, Petrovitch H. Current evidence for neuroprotective effects of nicotine and caffeine against Parkinson's disease. *Drugs Aging*. 2001;18(11):797-806.
43. Torti M, Vacca L, Stocchi F. Istradefylline for the treatment of Parkinson's disease: is it a promising strategy? *Expert Opin Pharmacother*. 2018;19(16):1821-1828.
44. Kostic VS, Svetel M, Sternic N, Dragasevic N, Przedborski S. Theophylline increases "on" time in advanced parkinsonian patients. *Neurology*. 1999;52(9):1916.
45. Ascherio A, Chen H, Schwarzschild MA, Zhang SM, Colditz GA, Speizer FE. Caffeine, postmenopausal estrogen, and risk of Parkinson's disease. *Neurology*. 2003;60(5):790-795.
46. Currie LJ, Harrison MB, Trugman JM, Bennett JP, Wooten GF. Postmenopausal estrogen use affects risk for Parkinson disease. *Arch Neurol*. 2004;61(6):886-888.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Maclagan LC, Visanji NP, Cheng Y, et al. Identifying drugs with disease-modifying potential in Parkinson's disease using artificial intelligence and pharmacoepidemiology. *Pharmacoepidemiol Drug Saf*. 2020; 1–9. https://doi.org/10.1002/pds.5015