

Use of medical care biases associations between Parkinson disease and other medical conditions

Anat Gross, MS, Brad A. Racette, MD, Alejandra Camacho-Soto, MD, Umber Dube, PhD, and Susan Searles Nielsen, PhD

Neurology® 2018;90:e2155-e2165. doi:10.1212/WNL.0000000000005678

Correspondence

Dr. Searles Nielsen
searles-nielsens@wustl.edu

Abstract

Objective

To examine how use of medical care biases the well-established associations between Parkinson disease (PD) and smoking, smoking-related cancers, and selected positively associated comorbidities.

Methods

We conducted a population-based, case-control study of 89,790 incident PD cases and 118,095 randomly selected controls, all Medicare beneficiaries aged 66 to 90 years. We ascertained PD and other medical conditions using ICD-9-CM codes from comprehensive claims data for the 5 years before PD diagnosis/reference. We used logistic regression to estimate age-, sex-, and race-adjusted odds ratios (ORs) between PD and each other medical condition of interest. We then examined the effect of also adjusting for selected geographic- or individual-level indicators of use of care.

Results

Models without adjustment for use of care and those that adjusted for geographic-level indicators produced similar ORs. However, adjustment for individual-level indicators consistently decreased ORs: Relative to ORs without adjustment for use of care, all ORs were between 8% and 58% lower, depending on the medical condition and the individual-level indicator of use of care added to the model. ORs decreased regardless of whether the established association is known to be positive or inverse. Most notably, smoking and smoking-related cancers were positively associated with PD without adjustment for use of care, but appropriately became inversely associated with PD with adjustment for use of care.

Conclusion

Use of care should be considered when evaluating associations between PD and other medical conditions to ensure that positive associations are not attributable to bias and that inverse associations are not masked.

CME Course

NPub.org/cmelist

From the Department of Neurology (A.G., B.A.R., A.C.-S., U.D., S.S.N.), Washington University School of Medicine, St. Louis, MO; and School of Public Health (B.A.R.), Faculty of Health Sciences, University of the Witwatersrand, Parktown, South Africa.

Go to Neurology.org/N for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

Glossary

AD = Alzheimer disease; **BASF** = beneficiary annual summary file; **CCI** = Charlson comorbidity index; **ICD-9-CM** = *International Classification of Diseases, Ninth Revision, Clinical Modification*; **OR** = odds ratio; **PCSA** = primary care service area; **PD** = Parkinson disease.

Epidemiologic studies examining the association between medical conditions and Parkinson disease (PD) may provide valuable insight into the etiology and pathophysiology of PD. However, access to and use of medical care may bias the association between 2 medical conditions,¹ but this bias rarely is considered in studies of PD. Even when all participants in a study have health insurance, their use of medical care, i.e., actual interactions with the health care system, could still bias the observed association between PD and another medical condition. The observed association can be distorted because people who obtain more medical care have more opportunity than others to be diagnosed with a variety of medical conditions.¹ We hypothesized that use of care would bias associations between PD and other medical conditions and, specifically, that the association would be biased upward. This is important because true inverse associations might be missed, while observed positive associations might be inflated and, thus, potentially spurious. Therefore, in a large, population-based study of universally insured patients, we examined how use of medical care biases several well-established associations between PD and other medical conditions.

Methods

Standard protocol approvals, registrations, and patient consents

This study required only deidentified data and was approved by the institutional review board at Washington University in St. Louis and by the Centers for Medicare and Medicaid Services.

Study design and eligibility criteria

We conducted a large, population-based, case-control study using detailed Medicare data as described previously.² Briefly, we used carrier (physician/supplier Part B), outpatient, inpatient, skilled nursing facility, durable medical equipment, and home health care claims from 2004 to 2009 and the 2009 beneficiary annual summary file (BASF) to construct the sample, obtain demographic data, and identify all medical conditions. All participants were age-eligible for Medicare by January 1, 2007 (born by February 1, 1942) and met the following criteria for 2009: Medicare Part A or B coverage, no Medicare Advantage plan coverage, residence in any of the 50 states in the United States, and age 90 years or younger at year end (or death in 2009 after diagnosis/reference).

PD case identification

We identified incident PD cases using ICD-9-CM diagnosis codes. We included eligible Medicare beneficiaries with at

least one ICD-9-CM code for PD (332 or 332.0) in 2009, but not in 2004–2008, and without Lewy body dementia (ICD-9-CM 331.82) or other neurodegenerative disease of the basal ganglia (ICD-9-CM 333 or 333.0).² In total, we included 89,790 incident PD cases, representing all 2009 beneficiaries with incident PD who met study eligibility criteria. This number of cases is consistent with prior studies of PD in Medicare.³

Control selection

We identified beneficiaries who had no ICD-9-CM code for PD or the similar conditions noted above in 2004–2009 but who otherwise met the same study eligibility criteria as cases. We randomly selected 118,095 as controls, representing 0.5% of all potential controls.²

Ascertainment of other medical conditions

We used ICD-9-CM codes in 2004–2009 prior to the PD diagnosis or a randomly assigned control reference date in 2009 to identify the presence/absence of selected medical conditions known as associated with PD. We included 6 conditions known to be positively associated^{4–8} with PD: REM sleep behavior disorder, anosmia (within a code for smell or taste disturbance), constipation, dementia/Alzheimer disease (AD),⁹ depression, and anxiety. Smoking is the most well-established factor inversely associated with PD,¹⁰ and cancer is generally inversely associated with PD, as well.⁴ Therefore, we evaluated tobacco smoking and the 2 cancers with very high (>80%) attributable risk for smoking (lung cancer and laryngeal cancer).^{11–13} We made a dichotomous ever vs never smoking variable with 4 codes specific to tobacco use: history of smoking (ICD-9-CM V15.82), nicotine dependence (ICD-9-CM 305.1), and 2 procedure codes for tobacco cessation counseling (99406 and 99407). We also made a more comprehensive ever smoking variable. This variable was a continuous variable that represented the probability of having ever smoked. We assigned all participants with a smoking-specific code or either of the above 2 smoking-related cancers a 100% probability of ever smoking. We assigned all other participants a probability that could range from 0% to 100%. Specifically, we used a logistic regression model based on >600 diagnosis and procedure codes indicative of 15 other conditions, along with demographic factors associated with smoking (sex, race/ethnicity, and birth cohort) to estimate this probability of ever smoking. We calculated and validated this probability as described previously.² For the present work, we retained this probability as a continuous variable that could range from 0 (never smoked) to 1 (ever smoked) and that also contained intermediate values. We coded the variable this way so that

the interpretation of the respective odds ratio (OR) would be comparable to the above dichotomous ever smoking variable.

Use of medical care variables

We obtained or created 5 continuous variables to quantify the use of medical care prior to PD diagnosis/reference. Two were geographic-level indicators: primary care use by Medicare beneficiaries in the region of residence and travel time from the region of residence to an urban area (where most neurologists and specialists practice). Three were individual-level indicators: number of visits (inpatient, outpatient, skilled nursing facility, physician office, and a dichotomous indicator for home health agency visits), number of unique diagnoses (5-digit ICD-9-CM codes), and number of different types of physicians/providers seen.

For the 2 geographic-level indicators, we linked the beneficiary's zip code of residence in 2009 from the BASF to publicly available data using the geographic information systems software ArcMap version 10.4.1 (Esri, Redlands, CA). For travel time to an urban area, we used a publicly available 5-digit zip code-level variable "TDIS1B" from U.S. Census-based data on rural-urban commuting area codes¹⁴ and then weighted this categorical variable by travel time to an urban center (0, 15, 37, 52, 67, 82, 90 minutes, i.e., each category's midpoint). For primary care use by Medicare beneficiaries in the region, we used the "PAR_PC10" variable from the US Health Resources & Service Administration data, which is the age-, sex-, and race-adjusted rate of total number of "primary care visits" in 2010 per Medicare beneficiary in the primary care service area (PCSA).¹⁵ A PCSA is the smallest geographic area that can be considered a discrete service area for primary medical care.¹⁶

For the individual-level indicators of use of care, we used the beneficiary's records in the 2004–2008 BASFs and detailed claims data from 2004–2009 up to diagnosis/reference. We did not use the 2009 BASF because it includes information after diagnosis/reference. First, across the linked 2004–2008 BASFs, we computed 5 "visit" variables. We summed each of the following: total number of days admitted as an inpatient ("IPSTY" variable), total number of days in a skilled nursing facility ("SNF_COVDYS"), total number of outpatient visits ("OPSTY"), and total number of physician/Part B carrier visits ("PHSVST"). We also computed a dichotomous indicator for home health agency visits. Second, we summed the total number of unique 5-digit ICD-9-CM codes observed in any of the 5 files detailed above. We did this for 2004–2008 for comparability to the visit variables, and then for completeness, created a second count variable including 2009 claims data up to diagnosis/reference. Third, we counted the total number of unique types of physicians/providers seen in the same period.

In addition, we calculated morbidity indices using all diagnosis codes up to diagnosis/reference. Specifically, we calculated a modified Charlson comorbidity index (CCI),¹⁷ while weighting categories using the Deyo¹⁸ and Quan¹⁹ methods.

We also calculated 28 dichotomous modified Elixhauser comorbidity measures²⁰ while excluding PD (332.0) and extrapyramidal diseases (333.9, 333.99). Although these indices are not measures of use of care per se, the CCI and Elixhauser index are validated indicators of mortality, which could affect use of medical care.

Statistical analysis

We performed all statistical analyses in Stata MP version 14.2 (StataCorp LP, College Station, TX). Our goal was to understand how bias from use of medical care distorts the association between PD and other medical conditions. We examined this through a 3-step approach. In the first step, we calculated the OR between PD and one other medical condition while adjusting for age, sex, and race, which are demographic factors known as associated with PD.³ In the second step, we calculated the OR between PD and the same medical condition while adjusting for these same demographic variables and for the specified use-of-care variable(s). Adjustment is a standard method for eliminating bias caused by a confounder. Although, technically, use of care can also contribute to the association between PD and another condition as a mediator, this contribution is not of interest because it does not reflect shared pathophysiology, genetics, or putative exposures. In this instance, adjustment for use of care produces ORs between PD and the medical condition above and beyond the contribution of use of care as a mediator while simultaneously removing the effect of use of care as a confounder. Finally, in the third step, we compared the above 2 age-, sex-, and race-adjusted ORs (with vs without adjustment for use of care). We formally compared these 2 ORs by calculating the percentage change (absolute difference in ORs divided by the OR without adjustment for use of care).

We calculated all ORs and their respective 95% confidence intervals through unconditional multivariable logistic regression models in which PD was the outcome variable and the other condition of interest was the "exposure" variable. We adjusted for age continuously as 2 linear splines with a knot at age 85. We adjusted for race in 7 categories: white, black, Asian, Pacific Islander/other, Native American, Hispanic ethnicity, and unknown race/ethnicity. Only 0.9% of cases and controls were in the unknown race category. Otherwise, all exposure and covariate data were known. We modeled each use-of-care variable as a continuous measure. The only exceptions were the dichotomous home health care variable and the Elixhauser index variables. For the latter, we first introduced all 28 variables simultaneously. We then only retained those significantly associated with PD after Bonferroni correction. Otherwise, if a model contained more than one use-of-care variable, we verified that all were positively associated with PD, i.e., that none became inversely associated with PD due to collinearity.

We repeated all analyses while dividing cases approximately in half according to certainty of PD diagnosis as detailed

previously,² and then we compared each case group to all controls. For the more certain case group, we required at least one ICD-9-CM code from a neurologist or ≥ 3 PD codes in 2009 ($n = 50,395$), while the less certain group had < 3 PD codes in 2009, none from a neurologist ($n = 39,305$).

Data availability

Because data for the present work originated from the Centers for Medicare and Medicaid Services, the authors are not authorized to make these data available.

Results

Characteristics of participants

Most participants (89% cases, 86% controls) were non-Hispanic white, and slightly more than half (51% cases, 57% controls) were women, indicating significant differences between PD cases and controls, consistent with prior studies⁴ (both $p < 0.001$) (table 1). We likewise confirmed that PD risk increased with age. In addition, PD cases demonstrated markedly greater use of medical care by all individual-level measures of use of care as compared to controls (all $p < 0.001$, after adjusting for age, sex, and race). However, PD cases and controls were very similar regarding number of primary care visits in the PCSA. PD cases could reach an urban area in slightly less time on average than controls.

Associations between PD and selected medical conditions

When only adjusting for age, sex, and race, we did not observe the expected inverse association between PD and smoking, lung cancer, and laryngeal cancer. However, we did observe significant positive associations as expected between PD and constipation, REM sleep behavior disorder, anosmia, depression, anxiety, and dementia/AD. Models also adjusting for geographic-level indicators of use of care produced very similar ORs when compared to the respective model that did not adjust for use of care (table 2).

In contrast, when we adjusted for individual-level indicators of use of care, ORs between PD and all conditions of interest were affected, sometimes markedly (table 2). ORs were universally lower after adjustment for use of care regardless of the individual-level indicator selected or whether the association with PD is known to be positive or negative (inverse). These adjustments were particularly notable for conditions known to be negatively associated with PD: Smoking and related conditions became inversely associated with PD only after adjustment for use of care.

Adjustment for some individual-level use-of-care variables lowered the ORs of interest more than others. The effect was more marked for all selected medical conditions when adjusting for the total number of unique ICD-9-CM codes in 2004–2008, than when adjusting for the total number of medical visits in the same years (table 2). ORs decreased

further when we adjusted for the total number of unique ICD-9-CM diagnosis codes in 2004–2009, i.e., also counted diagnoses that occurred in the months immediately before PD diagnosis/reference. We verified that this more aggressive adjustment was not driven by the inclusion of codes that might indicate PD per se (ICD-9-CM codes 332.1–333.99 and 15 additional PD-predictive codes previously classified as PD symptoms²). Excluding these codes from the count of diagnosis codes resulted in ORs that were within 3% of the ORs adjusted for the total count of diagnosis codes (not shown in tables). The total number of type of physicians seen in 2004–2009 was strongly associated with the total number of unique ICD-9-CM codes for both cases and controls (Spearman $\rho = 0.82$ – 0.90). Accordingly, addition of this variable to the model with adjustment for count of diagnosis codes did not markedly change ORs further (not shown), but when alone, did decrease ORs markedly for all conditions (table 2). Similarly, ever smoking probability was also associated with unique ICD-9-CM codes for both cases and controls (Spearman $\rho = 0.21$ – 0.23), and adjustment for this in addition to count of diagnosis codes also did not materially change the ORs, with the exception of the smoking-related conditions, which were attenuated as expected.

When we adjusted for the comorbidity indices, the effect on ORs again was universally downward (table 3). Specifically, Elixhauser-adjusted ORs were relatively similar to those adjusted for count of ICD-9-CM codes for most conditions considered. However, adjusting for CCI, which weighs heavily for cancer, was notably less aggressive for most conditions, with the exception of lung and laryngeal cancer. Furthermore, the Elixhauser-adjusted ORs for depression, anxiety, and dementia/AD were markedly lower (adjusted further) as compared to ORs adjusted for other use-of-care variables.

Overall, adjusting for individual-level use-of-care variables decreased ORs between 8% and 58%, depending on the medical condition and use-of-care measure (table 4). Adjustment for the comorbidity indices had even more variable effects on the ORs, ranging effectively from no change to a 65% decrease. We confirmed all of these results in the sensitivity analysis in which we divided PD cases according to certainty of diagnosis and then compared each PD case group to all controls. That is, adjustment for use of care decreased ORs by a relatively similar percentage for all medical conditions, regardless of PD case certainty (table 4).

Discussion

This study provides strong evidence that the association between PD and other medical conditions can be biased substantially if use of medical care is not taken into account. After adjustment for use of care, our ORs for the conditions of interest had the same direction and a more similar magnitude to those reported in previous literature. Most notably, smoking became inversely associated with PD, as expected. A

Table 1 Characteristics of PD cases and controls, Medicare 2009

	PD cases (n = 89,790)	Controls (n = 118,095)
Female, %	50.2	57.0
Race/ethnicity, %		
White	88.8	86.4
Black	6.0	7.5
Pacific Islander/other	1.0	1.5
Asian	1.7	2.2
Hispanic	2.1	1.9
Native American	0.3	0.4
Unknown	0.09	0.09
Home health agency visits, 2004–2008, %	39.2	32.3
Age, y, mean (SD)	79 (6.1)	76 (6.2)
Continuous individual-level indicators of use of care,^a mean (SD)		
Unique 5-digit ICD-9-CM codes in 2004–2008	88.6 (50.7)	59.9 (43.8)
Unique 5-digit ICD-9-CM codes in 2004–2009	100.9 (54.3)	65.6 (47.1)
Inpatient visits in 2004–2008, d ^b	2 (2.8)	1.1 (1.9)
Outpatient visits in 2004–2008, n	21.7 (25.1)	15 (20.3)
Skilled nursing facility visits in 2004–2008, d	12.5 (35.1)	3.9 (18.7)
Physician/Part B carrier visits in 2004–2008, n	43.1 (36)	31.1 (30.7)
No. of different types of physicians/providers in 2004–2009	19.4 (6.7)	15.1 (6.8)
Geographic-level indicators of use of care, mean (SD)		
Travel time from home zip code to urban area, min	12.8 (23.2)	13.5 (24.0)
Primary care use by medicare beneficiaries in region ^c	3.8 (0.7)	3.8 (0.7)

Abbreviations: ICD-9-CM = *International Classification of Diseases, Ninth Revision, Clinical Modification*; PD = Parkinson disease.

^a Before PD diagnosis/reference date.

^b Includes beneficiaries without any inpatient visits as 0 days from 2004–2008.

^c Age-, sex-, and race-adjusted rate of number of “primary care visits” by Medicare beneficiaries in the primary care service area in 2010.

recent methods-oriented study that focused on the effect of use of care on the relationship between 3 pairs of medical conditions (diabetes and depression, depression and weight loss, and hypertension and myocardial infarction) observed similar results.¹ Like our study, ORs decreased markedly with adjustment for the number of medical visits and/or CCI.

Our findings almost certainly extend beyond the chosen conditions of interest. We consistently observed that without adjustment for use of care, the association between PD and any medical condition may be biased upward. This has important implications for prior studies of PD and other medical conditions. Typically, prior studies have not adjusted for use of care. Without this adjustment, one might derive incorrect conclusions from the observed results. Specifically, when a positive association between PD and another condition is observed, the association may be partly or fully attributable to bias by use of care. It is even possible that there is an inverse

association between PD and the condition. There are numerous examples of studies that report a novel medical “risk factor” for PD. While the reported positive association may provide valuable information on the pathophysiology of PD, it is also possible that adjustment for use of care could eliminate or even reverse the association. In contrast, when an inverse association is observed between PD and another medical condition, the association is unlikely to be attributable to bias from use of care. Rather, adjustment for use of care would likely strengthen the inverse association, moving the observed relative risk estimate further below null. Finally, if the observed association is null, it is unlikely that there is a positive association, but one cannot rule out an inverse association. As such, when examining the relationship between PD and another medical condition, the association or lack thereof should be interpreted in light of the potential for bias by use of care. Understanding the true relationship between PD and other medical conditions is critical to prioritize the PD

Table 2 Selected medical conditions and risk of PD, with and without adjustment for use of medical care

	Age-, sex-, and race-adjusted ORs (95% CI), ^a Medicare 2009								Prior studies	
	Cases (n = 89,790), %	Controls (n = 118,095), %	No adjustment for use of care	Adjusted for geographic- level indicators of use of care ^c	Adjusted for individual-level indicators of use of care ^b				Study design	OR or relative risk (95% CI)
					No. of visits 2004–2008 ^d	No. of unique ICD-9-CM codes 2004–2008	No. of unique ICD-9-CM codes 2004–2009	No. of provider types 2004–2009		
Medical conditions established as inversely associated with PD^e										
Lung cancer	2.3	1.8	1.23 (1.16–1.31)	1.20 (1.12–1.28)	0.92 (0.86–0.98)	0.83 (0.78–0.89)	0.76 (0.71–0.81)	0.77 (0.72–0.82)	Meta- analysis	0.48 (0.39–0.60) ¹¹
Laryngeal cancer	0.35	0.32	0.97 (0.83–1.13)	1.00 (0.85–1.18)	0.74 (0.64–0.87)	0.66 (0.56–0.78)	0.61 (0.52–0.72)	0.63 (0.54–0.74)	Matched case-control ^g	0.35 (0.14–0.86) ¹²
Ever smoked (smoking-specific codes)^e	19	17	1.20 (1.17–1.22)	1.19 (1.16–1.22)	0.95 (0.92–0.97)	0.86 (0.84–0.88)	0.80 (0.78–0.82)	0.83 (0.81–0.85)	Meta- analysis	0.59 (0.54–0.63) ¹⁰
Ever smoked (predicted probability)^f	— ^f	— ^f	1.65 (1.60–1.71)	1.65 (1.59–1.71)	1.02 (0.98–1.05)	0.82 (0.79–0.85)	0.69 (0.67–0.72)	0.76 (0.74–0.79)		
Medical conditions established as positively associated with PD^e										
Constipation	34	18	2.18 (2.14–2.23)	2.17 (2.12–2.22)	1.71 (1.67–1.75)	1.52 (1.49–1.55)	1.39 (1.36–1.43)	1.49 (1.46–1.53)	Meta- analysis	2.27 (2.09–2.46) ⁵
REM sleep behavior disorder	0.23	0.02	9.87 (6.66–14.6)	10.21 (6.58–15.8)	9.08 (6.11–13.5)	7.76 (5.22–11.5)	7.48 (5.02–11.1)	6.74 (4.53–10.0)	Meta- analysis	3.54 (2.77–4.52) ⁶
Anosmia (or taste disturbance)	0.78	0.44	1.72 (1.53–1.94)	1.69 (1.50–1.92)	1.51 (1.34–1.71)	1.32 (1.17–1.49)	1.27 (1.12–1.43)	1.23 (1.10–1.39)	Prospective cohort ^h	5.20 (1.50–25.6) ⁷
Depression	44	20	3.37 (3.30–3.44)	3.38 (3.31–3.45)	2.70 (2.64–2.76)	2.49 (2.44–2.54)	2.28 (2.23–2.33)	2.39 (2.34–2.44)	Matched case-control ⁱ	1.9 (1.10–3.20) ⁸

Continued

Table 2 Selected medical conditions and risk of PD, with and without adjustment for use of medical care (continued)

	Age-, sex-, and race-adjusted ORs (95% CI), ^a Medicare 2009								Prior studies	
	Cases (n = 89,790), %	Controls (n = 118,095), %	No adjustment for use of care	Adjusted for geographic- level indicators of use of care ^c	Adjusted for individual-level indicators of use of care ^b					
					No. of visits 2004–2008 ^d	No. of unique ICD-9-CM codes 2004–2008	No. of unique ICD-9-CM codes 2004–2009	No. of provider types 2004–2009	Study design	OR or relative risk (95% CI)
Anxiety	32	17	2.39 (2.34–2.45)	2.39 (2.33–2.44)	1.93 (1.88–1.97)	1.75 (1.71–1.79)	1.63 (1.59–1.66)	1.74 (1.70–1.79)	Matched case-control ⁱ	2.2 (1.40–3.40) [§]
Dementia/AD	34	7.9	5.15 (5.02–5.29)	5.18 (5.04–5.33)	4.55 (4.43–4.68)	4.24 (4.13–4.35)	3.99 (3.88–4.09)	4.20 (4.09–4.31)		— ^j

Abbreviations: AD = Alzheimer disease; CI = confidence interval; ICD-9-CM = *International Classification of Diseases, Ninth Revision, Clinical Modification*; OR = odds ratio; PD = Parkinson disease.

^a OR and 95% CI, adjusted for sex, age (continuously using 2 linear splines with a knot at age 85), and race (as 7 categories: white, black, Asian, Pacific Islander/other, Native American, Hispanic ethnicity, and unknown race/ethnicity).

^b Before PD diagnosis/reference date. ICD-9-CM codes counted as 5-digit codes.

^c Age-, sex-, and race-adjusted rate of number of "primary care visits" by Medicare beneficiaries in the primary care service area in 2010 and the travel time to the closest urban area in 2010 from the beneficiary's 2009 residential zip code, both as continuous measures.

^d Total number of inpatient, outpatient, skilled nursing facility, physician/Part B carrier, as continuous variables, and a dichotomous variable for home health agency visits.

^e Medical conditions before PD diagnosis/reference date, ascertained using the following ICD-9-CM codes: lung cancer (162.x, 231.2, V10.11), laryngeal cancer (161.x, V10.21), ever smoking (V15.82, 305.1, and Current Procedural Terminology codes 99406 and 99407), REM sleep behavior disorder (327.42), anosmia (781.1, within a code for smell or taste disturbance), constipation (564.0x), dementia/AD (290.0, 290.1–290.13, 290.20, 290.21, 290.3, 290.40–290.43, 291.2, 294.10, 294.11, 046.1, 331.0, 331.11, 331.19, 331.2, 331.82),[§] depression (296.20–296.25, 296.30–296.35, 296.5, 296.6, 296.82, 296.90, 300.4, 309.0, 309.1, 309.28, and 311.0),²⁹ and anxiety (300.0–300.9).³⁰

^f Percent not applicable. The predicted probability of ever smoking is a continuous variable ranging from 0 to 1, estimated previously² from 17 conditions identified using >600 ICD-9-CM and procedure codes prior to PD diagnosis/reference date, sex, race, and birth cohort. These codes include the 4 smoking-specific codes and ICD-9-CM codes for lung cancer and laryngeal cancer specified in the above footnote.

[§] Participants in this study were matched by sex and year of birth.

^h The model adjusted for age, smoking, coffee consumption, constipation, excessive daytime sleepiness, and Cognitive Abilities Screening Instrument score.

ⁱ Participants in this study were matched by sex and age.

^j Studies found analyzed the association between dementia in family members and PD. These studies observed ORs between 2 and 3, which we used to check the reasonability of our results.⁴

Table 3 Selected medical conditions and risk of PD with adjustment for comorbidity indices

	Age-, sex-, and race-adjusted ORs (95% CI) ^a		
	Adjusted for Charlson comorbidity index ^b	Adjusted for all Elixhauser comorbidity variables ^c	Adjusted for selected Elixhauser comorbidity variables ^d
Medical conditions established as inversely associated with PD^e			
Lung cancer	0.64 (0.60–0.69)	0.91 (0.85–0.98)	0.90 (0.84–0.97)
Laryngeal cancer	0.59 (0.50–0.69)	0.75 (0.63–0.88)	0.74 (0.63–0.88)
Ever smoked (smoking-specific codes) ^e	0.92 (0.90–0.95)	0.84 (0.82–0.87)	0.85 (0.83–0.87)
Ever smoked (predicted probability) ^f	0.91 (0.88–0.94)	0.81 (0.78–0.84)	0.83 (0.80–0.86)
Medical conditions established as positively associated with PD^e			
Constipation	1.81 (1.77–1.85)	1.50 (1.47–1.54)	1.41 (1.38–1.45)
REM sleep behavior disorder	9.99 (6.72–14.9)	8.01 (5.35–12.0)	7.44 (4.95–11.2)
Anosmia (or taste disturbance)	1.62 (1.44–1.83)	1.46 (1.29–1.66)	1.39 (1.22–1.57)
Depression	2.77 (2.71–2.83)	1.96 (1.91–2.00)	1.32 (1.23–1.41)
Anxiety	2.04 (2.00–2.09)	1.57 (1.53–1.61)	1.28 (1.25–1.32)
Dementia/AD	4.17 (4.06–4.28)	2.76 (2.68–2.84)	1.80 (1.75–1.86)

Abbreviations: AD = Alzheimer disease; CI = confidence interval; ICD-9-CM = *International Classification of Diseases, Ninth Revision, Clinical Modification*; OR = odds ratio; PD = Parkinson disease.

^a OR and 95% CI, based on 89,700 cases and 118,095 controls, adjusted for sex, age (continuously using 2 linear splines with a knot at age 85), race (as 7 categories: white, black, Asian, Pacific Islander/other, Native American, Hispanic ethnicity, and unknown race/ethnicity), and the specified comorbidity index.

^b A single continuous score ranging from 0 to 33 based on 17 weighted dichotomous indicators of myocardial infarction, congestive heart failure, peripheral vascular disease, cerebrovascular disease, dementia, chronic pulmonary disease, mild liver disease, moderate or severe liver disease, diabetes, diabetes with chronic complications, renal disease, rheumatologic disease, peptic ulcer disease, hemiplegia, any malignancy, metastatic solid tumor, and AIDS.

^c Twenty-eight dichotomous Elixhauser comorbidity measures, excluding ICD-9-CM codes for PD (332.0) and extrapyramidal diseases indicative of PD (333.9, 333.99).

^d Seventeen dichotomous Elixhauser comorbidity measures significantly associated with PD including variables for congestive heart failure, pulmonary circulation disorders, peripheral vascular disease, hypertension, paralysis, other neurologic disorders, hypothyroidism, metastatic cancer, rheumatoid arthritis, obesity, weight loss, fluid electrolyte disorders, deficiency anemias, psychoses, and depression.

^e Medical conditions before PD diagnosis/reference date, ascertained using the following ICD-9-CM codes: lung cancer (162.x, 231.2, V10.11), laryngeal cancer (161.x, V10.21), ever smoking (V15.82, 305.1, and Current Procedural Terminology codes 99406 and 99407), REM sleep behavior disorder (327.42), anosmia (781.1, within a code for smell or taste disturbance), constipation (564.0x), dementia/AD (290.0, 290.1–290.13, 290.20, 290.21, 290.3, 290.40–290.43, 291.2, 294.10, 294.11, 046.1, 331.0, 331.11, 331.19, 331.2, 331.82),⁹ depression (296.20–296.25, 296.30–296.35, 296.5, 296.6, 296.82, 296.90, 300.4, 309.0, 309.1, 309.28, and 311.0),²⁹ and anxiety (300.0–300.9).³⁰

^f Continuous variable ranging from 0 to 1, estimated previously² from 17 conditions identified using >600 ICD-9-CM and procedure codes prior to PD diagnosis/reference date, sex, race, and birth cohort.

research agenda as well as to guide the development of potential biomarkers and therapeutics.

Our analysis also provides practical information for future analyses of the relation between PD and other medical conditions. As evidenced by our study, restriction to a population with access to health insurance is not sufficient. Our statistical approach is easy to implement and is appropriate if use of care acts as a confounder or mediator. Comparing logistic regression models with and without adjustment for use of care, we explored the extent to which several indicators addressed bias by use of care. This approach is generally appropriate, and any potential bias that could be introduced would be negligible when the 2 diseases of interest are not perfectly ascertained by the methods of ascertainment.²¹ Of note, for all conditions of interest, the individual-level indicators of use of care led to marked corrections to the ORs. In addition, measures that captured the number of diagnoses, number of

types of physicians seen, and the Elixhauser comorbidities, as opposed to the number of visits and CCI, appeared to more aggressively adjust ORs of PD and most conditions of interest. Similarly, individual-level indicators derived from claims data up to PD diagnosis/reference decreased ORs more than indicators restricted to the years prior. The months immediately before PD diagnosis may legitimately contribute to greater ascertainment of PD. Our prior research in this dataset,²² and research conducted in other large case-control studies, indicates that the positive association between PD and various traumas is greatest in the 3 months before PD diagnosis.^{23,24}

In practice, potentially only one individual-level variable is sufficient for adjustment for use of care. The number of diagnosis codes and number of types of physicians seen are relatively simple to calculate, and adjusting for them brought the well-established PD-smoking association closest to the

Table 4 Percentage decrease in the association^a between selected medical conditions^b and PD when adjusting for use of medical care or comorbidity indices, overall and by PD case certainty^c

	Individual-level indicators of use of medical care ^d				Comorbidity indices	
	No. of visits 2004–2008	No. of unique ICD-9-CM codes 2004–2008	No. of unique ICD-9-CM codes 2004–2009	No. of provider types 2004–2009	Charlson comorbidity index	Selected Elixhauser comorbidity variables
All PD cases vs controls^c						
Lung cancer	25.2	32.5	38.2	37.4	48.0	26.8
Laryngeal cancer	22.7	32.0	37.1	35.0	39.2	23.7
Ever smoked (smoking codes) ^b	20.8	28.3	33.3	30.8	23.3	29.2
Ever smoked (probability) ^f	38.2	50.3	58.2	53.9	44.8	49.7
Constipation	21.6	30.3	36.2	31.6	16.9	35.3
REM sleep behavior disorder	8.0	21.4	24.2	31.7	— ^e	24.6
Anosmia (or taste disturbance)	12.2	23.2	26.2	28.5	5.8	19.2
Depression	19.9	26.1	32.3	29.1	17.8	60.8
Anxiety	19.2	26.8	31.8	27.2	14.6	46.4
Dementia/AD	11.6	17.7	22.5	18.4	19.0	65.0
More certain PD cases vs controls^c						
Lung cancer	24.3	31.8	36.4	39.2	42.1	21.4
Laryngeal cancer	22.6	30.1	34.4	35.5	34.4	17.2
Ever smoked (smoking codes) ^b	19.3	27.5	31.2	32.1	20.2	26.6
Ever smoked (probability) ^f	36.9	49.6	56.0	56.0	40.4	46.1
Constipation	19.9	28.0	33.0	32.6	14.5	28.0
REM sleep behavior disorder	9.2	20.1	22.7	31.5	— ^e	21.7
Anosmia (or taste disturbance)	13.3	22.3	24.5	29.3	5.3	15.4
Depression	17.5	24.8	29.8	30.8	14.6	38.1
Anxiety	17.8	25.2	29.6	27.8	12.2	31.7
Dementia/AD	11.2	18.8	23.0	21.4	17.0	45.3
Less certain PD cases vs controls^c						
Lung cancer	25.3	33.1	39.4	35.2	52.1	26.8
Laryngeal cancer	22.8	33.7	39.6	32.7	44.5	25.7
Ever smoked (smoking codes) ^b	21.2	28.0	34.1	28.0	24.2	30.3
Ever smoked (probability) ^f	37.7	49.7	58.1	49.2	47.6	51.8
Constipation	23.3	32.9	39.5	30.0	20.9	35.7
REM sleep behavior disorder	8.1	26.1	30.8	32.3	6.7	23.1
Anosmia (or taste disturbance)	14.1	28.2	32.2	30.2	10.1	22.1
Depression	21.5	27.1	34.6	26.8	20.4	46.4

Continued

Table 4 Percentage decrease in the association^a between selected medical conditions^b and PD when adjusting for use of medical care or comorbidity indices, overall and by PD case certainty^c (continued)

	Individual-level indicators of use of medical care ^d				Comorbidity indices	
	No. of visits 2004–2008	No. of unique ICD-9-CM codes 2004–2008	No. of unique ICD-9-CM codes 2004–2009	No. of provider types 2004–2009	Charlson comorbidity index	Selected Elixhauser comorbidity variables
Anxiety	20.9	28.3	34.4	25.4	16.8	38.1
Dementia/AD	12.9	18.5	23.8	17.5	21.0	48.6

Abbreviations: AD = Alzheimer disease; ICD-9-CM = *International Classification of Diseases, Ninth Revision, Clinical Modification*; PD = Parkinson disease.

^a Odds ratio adjusted for age, sex, race, and specified indicator of use of medical care or comorbidity index vs the respective odds ratio adjusted only for age, sex, and race.

^b Medical conditions before PD diagnosis/reference date, ascertained using the following ICD-9-CM codes: lung cancer (162.x, 231.2, V10.11), laryngeal cancer (161.x, V10.21), ever smoking (V15.82, 305.1, and Current Procedural Terminology codes 99406 and 99407), REM sleep behavior disorder (327.42), anosmia (781.1, within a code for smell or taste disturbance), constipation (564.0x), dementia/AD (290.0, 290.1–290.13, 290.20, 290.21, 290.3, 290.40–290.43, 291.2, 294.10, 294.11, 046.1, 331.0, 331.11, 331.19, 331.2, 331.82),⁹ depression (296.20–296.25, 296.30–296.35, 296.5, 296.6, 296.82, 296.90, 300.4, 309.0, 309.1, 309.28, and 311.0),²⁹ and anxiety (300.0–300.9).³⁰

^c More certain PD cases (n = 50,395) are those with an ICD-9-CM code for PD from a neurologist or ≥3 ICD-9-CM codes for PD in 2009. Less certain PD cases (n = 39,305) are those with <3 ICD-9-CM codes for PD in 2009, none from a neurologist. All ORs were based on the specified PD case group relative to all controls (n = 118,095).

^d Before PD diagnosis/reference date. Visits are the total number of inpatient, outpatient, skilled nursing facility, physician/Part B carrier, as continuous variables, and a dichotomous variable for home health agency visits. Counts of unique ICD-9-CM diagnosis codes are based on 5-digit codes.

^e The age-, sex-, and race-adjusted odds ratio did not materially change with adjustment for the Charlson comorbidity index (increased by <1%).

^f Continuous variable ranging from 0 to 1, estimated previously² from 17 conditions identified using >600 ICD-9-CM and procedure codes prior to PD diagnosis/reference date, sex, race, and birth cohort.

expected OR. Of these 2 single variables, the number of diagnosis codes decreased ORs more aggressively than did the number of provider types. Excluding codes indicative of PD from the total count of diagnosis codes might be appropriate to avoid overadjustment, but this subjective refinement affected results very minimally. In contrast, Elixhauser and CCI adjustments require multiple variables, and there might be greater potential for overadjustment for conditions that particularly influence these indices, although underadjustment has also been reported.²⁵ We observed evidence of both over- and underadjustment when adjusting for comorbidity indices instead of number of diagnosis codes. The potential for underadjustment is consistent with another Medicare-based study that found that the number of diagnoses was more predictive of 30-day mortality than the CCI.²⁶ And, finally, while the number of visit variables are conveniently already calculated in the Medicare base file, they did not adjust for use of care as aggressively as the other individual-level indicators, even when all used in combination. In summary, adjustment for some individual-level indicators of use of care is necessary when evaluating associations between medical conditions and PD, and there are some advantages to using the total unique count of diagnosis codes in particular. In contrast, geographic-level indicators are not sufficient substitutes for an individual-level indicator of use of care. This is unsurprising since we expect more error in geographic-level indicators. In addition, they were not strongly associated with PD in our study and, therefore, had limited ability to alter ORs.

The strengths of this study are the large population-based sample that included all eligible incident PD cases, a randomly selected and highly comparable control group, and a comprehensive set

of claims data. These data covered claims beyond inpatient and outpatient files, from which we were able to compute a wider range of measures of use of care. The main limitation of our study is that we cannot determine whether bias by use of care would have such marked effects in all studies of PD, or whether this effect is specific to Medicare data. It is possible that administrative claims data may be more affected by bias by use of care.²⁵ We may have been able to observe this phenomenon clearly in our study for several additional reasons. First, our claims data extended back 5 years prior to PD diagnosis, and it is generally accepted that the prodromal period extends back fully into this period.²⁷ Most PD cases were likely already symptomatic and, therefore, accelerating their interaction with the health care system. Second, because this was a population-based study, we were not forced to make controls overly similar to PD cases regarding use of medical care. Third, we restricted our study to age-eligible Medicare beneficiaries to make it a population-based sample. We cannot rule out the possibility that bias by use of care is accentuated in this relatively old (66–90 years) population-based sample of beneficiaries who only have Medicare coverage. Another limitation is that our dichotomous smoking variable was based on only a few diagnosis and procedure codes. The prevalence of smoking in both cases and controls is much lower than expected,²⁸ consistent with substantial misclassification that could account for our observed OR for smoking and PD being closer to the null than reported in previous studies. However, we observed a similar effect of adjustment for use of care when we used a more comprehensive smoking variable that yielded ORs much more similar to those observed in previous studies.¹⁰ Lastly, the exact magnitude of the ORs for the other conditions we considered is less well established. While this

limits our ability to interpret our results, we note the consistency of results across all conditions examined, as well as with the prior literature on the effect of adjusting for use of care.

Overall, our study findings suggest that adjustment for use of care should be considered when evaluating the association between PD and another medical condition. Our findings also suggest that some prior positive associations between novel risk factors and PD reported in the literature may be inflated and, therefore, perhaps spurious.

Author contributions

Anat Gross performed data analysis and coauthored the first draft of this manuscript. Brad A. Racette obtained funding for this study, oversaw data analysis, and edited this manuscript. Dr. Racette takes full responsibility for the data, the analyses and interpretation, and the conduct of the research; he has full access to all of the data; and he has the right to publish any and all data separate and apart from any sponsor. Alejandra Camacho-Soto assisted with creating variables and edited this manuscript. Umber Dube assisted with creating variables and edited this manuscript. Susan Searles Nielsen performed data analysis and coauthored the first draft of this manuscript.

Study funding

Study funded by the Michael J. Fox Foundation, National Institute of Environmental Health Sciences (K24ES017765), and American Parkinson Disease Association.

Disclosure

A. Gross reports no disclosures relevant to the manuscript. B. Racette reports research support from Teva (principal investigator [PI]), US WorldMeds (PI), Allergan (PI), and Vaccinex (PI); government research support from NIH (K24ES017765 [PI], R21ES17504 [PI], R01ES021488 [PI], R01ES021488-02S1 [PI], R01ES025991 [PI]); research support from the Michael J. Fox Foundation; and consultation income from 86 Pillars for legal consultation. A. Camacho-Soto reports government research support from NIH (K12HD00109719 [PI]). U. Dube reports no disclosures relevant to the manuscript. S. Searles Nielsen reports government research support from NIH (R21ES17504 [coinvestigator, co-I], R01ES021488 [co-I], R01ES021488-02S1 [co-I], R01ES025991 [co-I], R01ES026891 [co-I]) and research support from the Michael J. Fox Foundation. Go to Neurology.org/N for full disclosures.

Received August 31, 2017. Accepted in final form March 30, 2018.

References

1. Goldstein BA, Bhavsar NA, Phelan M, Pencina MJ. Controlling for informed presence bias due to the number of health encounters in an electronic health record. *Am J Epidemiol* 2016;184:847–855.

2. Searles Nielsen S, Warden MN, Camacho-Soto A, Willis AW, Wright BA, Racette BA. A predictive model to identify Parkinson disease from administrative claims data. *Neurology* 2017;89:1448–1456.
3. Wright Willis A, Evanoff BA, Lian M, Criswell SR, Racette BA. Geographic and ethnic variation in Parkinson disease: a population-based study of US Medicare beneficiaries. *Neuroepidemiology* 2010;34:143–151.
4. Wirdefeldt K, Adami HO, Cole P, Trichopoulos D, Mandel J. Epidemiology and etiology of Parkinson's disease: a review of the evidence. *Eur J Epidemiol* 2011;26(suppl 1):S1–S58.
5. Adams-Carr KL, Bestwick JP, Shribman S, Lees A, Schrag A, Noyce AJ. Constipation preceding Parkinson's disease: a systematic review and meta-analysis. *J Neurol Neurosurg Psychiatry* 2016;87:710–716.
6. Zhang J, Xu CY, Liu J. Meta-analysis on the prevalence of REM sleep behavior disorder symptoms in Parkinson's disease. *BMC Neurol* 2017;17:23.
7. Ross GW, Petrovitch H, Abbott RD, et al. Association of olfactory dysfunction with risk for future Parkinson's disease. *Ann Neurol* 2008;63:167–173.
8. Shiba M, Bower JH, Maraganore DM, et al. Anxiety disorders and depressive disorders preceding Parkinson's disease: a case-control study. *Mov Disord* 2000;15:669–677.
9. St Germaine-Smith C, Metcalfe A, Pringsheim T, et al. Recommendations for optimal ICD codes to study neurologic conditions: a systematic review. *Neurology* 2012;79:1049–1055.
10. Hernan MA, Takkouche B, Caamano-Isorna F, Gestal-Otero JJ. A meta-analysis of coffee drinking, cigarette smoking, and the risk of Parkinson's disease. *Ann Neurol* 2002;52:276–284.
11. Xie X, Luo X, Xie M, Liu Y, Wu T. Risk of lung cancer in Parkinson's disease. *Oncotarget* 2016;7:77319–77325.
12. Olsen JH, Friis S, Frederiksen K. Malignant melanoma and other types of cancer preceding Parkinson disease. *Epidemiology* 2006;17:582–587.
13. Whitman DC, Wilson LF. The fractions of cancer attributable to modifiable factors: a global review. *Cancer Epidemiol* 2016;44:203–221.
14. University of Washington Rural Health Research Center. Rural-urban commuting area (RUCA) zip code data, version 2.0 [online]. Available at: depts.washington.edu/uwruca/. Accessed July 18, 2016.
15. U.S. Department of Health and Human Services. Primary care service area (PCSA) data for 2010 U.S. Census tracts [online]. Available at: datawarehouse.hrsa.gov/data/datadownload/pcsadata2010download.aspx. Accessed July 15, 2016.
16. Goodman DC, Mick SS, Bott D, et al. Primary care service areas: a new tool for the evaluation of primary care services. *Health Serv Res* 2003;38:287–309.
17. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373–383.
18. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 1992;45:613–619.
19. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 2005;43:1130–1139.
20. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care* 1998;36:8–27.
21. Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology* 2003;14:300–306.
22. Camacho-Soto A, Warden MN, Searles Nielsen S, et al. Traumatic brain injury in the prodromal period of Parkinson's disease: a large epidemiological study using Medicare data. *Ann Neurol* 2017;82:744–754.
23. Nystrom H, Nordstrom A, Nordstrom P. Risk of injurious fall and hip fracture up to 26 y before the diagnosis of Parkinson disease: nested case-control studies in a nationwide cohort. *PLoS Med* 2016;13:e1001954.
24. Kenborg L, Rugbjerg K, Lee PC, et al. Head injury and risk for Parkinson disease: results from a Danish case-control study. *Neurology* 2015;84:1098–1103.
25. Schneeweiss S, Maclure M. Use of comorbidity scores for control of confounding in studies using administrative databases. *Int J Epidemiol* 2000;29:891–898.
26. Melfi C, Holleman E, Arthur D, Katz B. Selecting a patient characteristics index for the prediction of medical outcomes using administrative claims data. *J Clin Epidemiol* 1995;48:917–926.
27. Postuma RB, Berg D. Advances in markers of prodromal Parkinson disease. *Nat Rev Neurol* 2016;12:622–634.
28. Searles Nielsen S, Franklin GM, Longstreth WT, Swanson PD, Checkoway H. Nicotine from edible Solanaceae and risk of Parkinson disease. *Ann Neurol* 2013;74:472–477.
29. Fiest KM, Jette N, Quan H, et al. Systematic review and assessment of validated case definitions for depression in administrative data. *BMC Psychiatry* 2014;14:289.
30. Kisely S, Lin E, Gilbert C, Smith M, Campbell LA, Vasiladias HM. Use of administrative data for the surveillance of mood and anxiety disorders. *Aust N Z J Psychiatry* 2009;43:1118–1125.