Enhancing Case Ascertainment of Parkinson's Disease Using Medicare Claims Data in a Population-Based Cohort: The Cardiovascular Health Study

TG Ton[1], **ML Biggs**[2], **D Comer**[4], **L Curtis**[5], **S Hu**[1], **EL Thacker**[6], **S Searles Nielsen**[7], **JA Delaney**[3], **D Landsittel**[4], **WT Longstreth**[1,3], **H Checkoway**[7], and **S Jain**[8]

[1]Department of Neurology, School of Medicine, University of Washington, Seattle, WA

[2]Department of Biostatistics, School of Public Health, University of Washington, Seattle, WA

[3]Department of Epidemiology, School of Public Health, University of Washington, Seattle, WA

[4]Center for Research on Health Care, University of Pittsburgh, Pittsburgh, PA

[5]Duke Clinical Research Institute, Department of Medicine, Duke University School of Medicine, Durham, NC

[6]Department of Health Sciences, College of Life Sciences, Brigham Young University, Provo, UT

[7]Department of Family and Preventive Medicine, University of California – San Diego, La Jolla, CA.

[8]Department of Neurology, University of Pittsburgh School of Medicine, Pittsburgh, PA

Abstract

Purpose: We sought to improve a previous algorithm to ascertain Parkinson's disease (PD) in the Cardiovascular Health Study (CHS) by incorporating additional data from Medicare outpatient claims. We compared our results to the previous algorithm in terms of baseline prevalence and incidence of PD, as well as associations with baseline smoking characteristics.

Methods: Our original case ascertainment used self-reported diagnosis, anti-parkinsonian medication, and hospitalization discharge ICD-9 code. In this study, we incorporated additional data from fee-for-service Medicare claims, extended follow-up time, review of medical records, and adjudicated cause of death. Two movement disorders specialists adjudicated final PD status. We used logistic regression models and controlled for age, sex, and African American race.

Results: We identified 75 additional cases, but reclassified 80 previously identified cases as not having PD. We observed significant inverse association with smoking status (odds ratio=0.42; 95% confidence interval=0.22, 0.79), and inverse linear trends with pack-years (p=0.005), and cigarettes per day (p=0.019) with incident PD. All estimates were stronger than those from the previous algorithm.

Corresponding Author: Thanh G.N. Ton, PhD, MPH Department of Neurology, Box 359775 University of Washington Seattle, WA 98195-9775 Phone: 206-744-3251 Fax: 206-744-8787 thanhton@uw.edu.

Conclusions: Our enhanced method did not alter prevalence and incidence estimates compared to our previous algorithm. However, our enhanced method provided stronger estimates of association, potentially due to reduced level of disease misclassification.

INTRODUCTION

For conditions such as Parkinson's disease (PD), large population-based prospective cohort studies designed to examine specific outcomes may provide cost-efficient alternatives to new case-control or cohort studies. The use of existing cohorts to address important aspects of PD has resulted in the publication of numerous significant findings.[1-11] Similarly, several ongoing population-based cohorts with well characterized participants such as the Framingham Heart Study,[12] the Atherosclerosis Risk in Communities (ARIC) Study,[13] the Multiethnic Study of Atherosclerosis (MESA),[14] and the Cardiovascular Health Study (CHS)[15] provide long-term follow-up and an abundance of measured covariates, which can be used to address efficiently current hypotheses and generate new ones for outcomes other than cardiovascular disease, including PD. However, most cardiovascular cohorts, including CHS, have not implemented rigorous systematic screening for PD. In our previous work, we relied on self-report, use of anti-parkinsonian medication, and ICD-9 discharge codes from hospitalizations to identify participants with PD.[16] Due to probable misclassification of PD status, our reported results may represent attenuated estimates of the true underlying associations.[17] Improving PD case ascertainment in the CHS would reduce misclassification, result in more accurate findings, and potentially increase statistical power.

Linkage to administrative data such as those from the Centers for Medicare & Medicaid Services (CMS) may provide additional information to enhance case ascertainment including exact dates of services, amount of PD-related services, and whether an encounter with a neurologist occurred. The CMS administers the eligibility list for Medicare, the primary health insurer for nearly all of the U.S. population aged 65 and older.[18] All Medicare beneficiaries receive Part A benefits that include coverage for inpatient, home health, and hospice care services. Medicare beneficiaries may also enroll in Part B of the program, which covers services provided by other institutional and non-institutional providers. Strengths of Medicare claims data include exact dates of service, amount of PD-related services in both inpatient and outpatient settings, and whether an encounter with a neurologist occurred. A recent linkage of CMS with the CHS has allowed us to potentially enhance our PD case ascertainment method with Medicare claims. As part of a larger study to examine clinical predictors of PD in the CHS, we sought to improve our classification of PD by updating our case ascertainment. We compared the performance of our enhanced PD classification with the original algorithm[16] in terms of baseline prevalence, incidence, and estimates of association for demographic and baseline smoking characteristics. We also describe the relative contribution of CMS to the classification of PD and its effect on corresponding results.

METHODS

The CHS is a prospective cohort study of coronary heart disease and stroke in adults aged 65 years and older initiated in 1989. The CHS recruited a total of 5,888 elderly men and women from four communities in the USA. Details about recruitment are described elsewhere [15]. In brief, 5,201 men and women were recruited into the original cohort, and an additional 687

African Americans were recruited three years after the initial baseline survey. Extensive physical and laboratory evaluations were performed at baseline to identify the presence and severity of cardiovascular risk factors. Participants were queried twice per year for new diagnoses, hospitalizations and medical procedures, and examined annually from 1989 to 1999, after which time twice-yearly telephone follow-up was continued. Reports of hospitalizations were cross-examined in the CMS records to enhance surveillance of events and deaths.

We provided CMS with the Social Security Number, sex, and date of birth for each of the 5888 CHS participants. Of those, 4549 (77.3%) were linked to Medicare enrollment records and Medicare Part A and Part B claims from January 1991 through December 2006 were provided. In order to comply with CMS's data use agreement that no mathematical formulas may be used if they result in the display of a cell 10 or less, we have slightly modified exposure categories such that all cells displayed and analyzed are 11. As such, current smokers also included former smokers who had quit within the past year. Likewise, we slightly modified categories for cigarettes per day to include: never smoker, <10, 10-19, 20-28, 29.

We detailed our original method of PD case ascertainment in a previous report.[16] Briefly, between 1989 and 2002, we identified participants who self-reported a physician diagnosis of PD; who used anti-parkinsonian medications reported at study visits; and who had hospitalization discharge records that contained the ICD-9 code 332.0. We identified 60 participants with prevalent PD defined as PD at baseline or identified during first year of follow-up, and 154 incident PD over the remaining years. In the current study, we enhanced our PD definition by: (1) extending follow-up an additional five years; (2) enhancing the screening process using data from CMS FFS; (3) including adjudicated cause-of-deaths as a data source; (4) reviewing medical records of participants who screened positive for potential PD; and (5) implementing a formal adjudication process to determine the final PD status. These enhancements are described in an online supplement, which details the sources of data used to screen for and adjudicate PD. Figure 1 provides an overview of data sources used in the CHS and CMS to identify PD. A comparison of the data sources used in the original algorithm and those used in the current enhanced ascertainment method is provided in online supplement (online Table A).

We used kappa statistics to compare the agreement between our two adjudicators before they reached a consensus. We defined PD as a dichotomous variable with those adjudicated as "probable PD" constituting the "yes" category and those adjudicated as either "possible PD" or "no PD" constituting the "no" category. As in our previous work, we defined prevalent PD as those with PD at study baseline or within the first year of follow-up. The remaining subjects identified with PD during follow-up were defined as incident. We calculated the age- and sex-specific point prevalence of PD at baseline and incidence over the remaining duration of follow-up.

We used logistic regression models to estimate the strength of associations between smoking characteristics and PD, and to allow us to make direct comparisons with results from our previous work.[16] Smoking characteristics represented our primary exposures of interest because an inverse association between smoking and PD is well established.[19] As a secondary analysis, we assessed potential confounding by time eligible in CMS. Because the total time a participant is eligible and enrolled in Medicare is correlated with the number of claims that can exist, time eligible in CMS could potentially be associated with the outcome, in this case, PD. As a potential confounder, time eligible and enrolled in CMS could also be related to risk factors of interest, such as smoking behavior as a proxy for lifestyle choices. For instance, a study assessing the association between smoking behavior and use of

primary care services among 254,382 adults 45 years and older with universal healthcare in Australia found that current smokers were less likely than non-smokers to make a Medicare claim [20]. The direction of such a systematic behavior among smokers would bias the estimates downward, potentially artificially producing or at least strengthening inverse associations with current smoking. As a result, we assessed potential confounding by time eligible in CMS by comparing estimates from models that included or excluded time eligible in CMS as a covariate.

Because we expected the CMS FFS claims data to improve our ability to pinpoint an "onset date", we also conducted a time-to-event analysis using Cox regression models. We defined "date of first evidence" of PD as the earliest time when PD appeared in any of the data sources. We defined loss to follow-up as the last date of any available information from CHS or CMS FFS. Participants were censored at time of death or at time of loss to follow-up, whichever came first. Because we defined prevalent PD as occurring at baseline and within the first year of follow-up, time at risk was calculated starting 365.25 days after enrollment until date of first evidence of PD, loss to follow-up, death, or end of follow-up, whichever occurred the earliest.

Finally, to isolate the contribution of the CMS FFS claims data to the ascertainment of PD, we restricted our analysis to the conditions of the original algorithm in which follow-up was only through 2002 rather than 2006; in which no information was available from adjudicated deaths and no medical records from hospitalizations were reviewed; and for which no adjudication was implemented for PD. We conducted logistic regression to compare associations for smoking when PD was defined in CHS, in CMS FFS, in either CHS or CMS FFS, and in both CHS and CMS FFS.

RESULTS

Using the enhanced case identification method, we identified a total of 463 of 5888 (7.9%) participants as potentially having PD or developing PD during follow-up, for whom adjudication for PD was necessary. Among these 463, evidence of PD was noted by adjudicated causes of death in 15.1%, by self-report in 35.0%, by medications in 39.5%, by ICD-9 inpatient diagnosis codes in 43.6%, and in CMS FFS claims in 27.2%. Medical records were available for 441 (95%) of 463 with some evidence of PD. In reviewing these medical records, 161 (36.5%) contained evidence supporting a diagnosis of PD while 20 (4.5%) contained evidence refuting PD. The remaining 260 (59.0%) of medical records had no mention of PD. While drug-induced parkinsonism was initially excluded based on the timing of a potential drug and the incident date, an additional 52 participants were discovered to have drug-induced parkinsonism during medical record review.

During the adjudication process, the two independent movement disorders specialists achieved 87.3% overall agreement (kappa = 0.80). Agreement between the two reviewers was greatest for participants with evidence of PD in both CHS and CMS FFS (kappa = 0.87); the lowest agreement occurred for those whose PD was identified only in CHS data (kappa = 0.65). After consensus was reached for all conflicting accounts, the final adjudication resulted in 209 (45.1%) with "Probable PD"; 71 (15.3%) with "Possible PD"; and 183 (39.5%) with "No PD." For subsequent analyses, the 209 were considered to have PD. Among these, 44 (21.1%) had prevalent disease, and the remaining 165 (78.9%) were classified as incident PD. Concerning 260 medical records that did not mention PD, 57 (21.9%) eventually had a final adjudicated status of "probable PD" and an additional 62(23.9%) were eventually considered "possible PD".

Based on the 44 subjects with prevalent PD, we estimated an overall prevalence of 0.7% (95% CI: 0.5%, 1.0%). Increasing age was associated with increasing prevalence (p=0.11). We also observed higher PD prevalence among men (p=0.004), among those with more than a high school degree (p=0.02), and among non-African Americans (p=0.014). A total of 165 developed PD over 70,796 person-years of follow-up, for an incidence rate of 233.1 per 100,000 person-years (95% CI: 200.1, 271.5). Standardized to the 1990 US Population, this rate is equivalent to 42.3 per 100,000. Rates were higher among men, and increased with older age (Table 1).

In comparing our results to the original method, the enhanced method identified an additional 75 participants as having PD, but also reclassified 80 participants previously considered as having PD to a final "not PD" status (Online Table C). On average, the new incident date occurred 4.3 years earlier than the old incident date.

Results for multivariate models using logistic regression are presented in Table 2 alongside results from our previous study using the original algorithm. In final models, we consistently observed statistically significant inverse linear trends for almost all baseline smoking characteristics. Similarly, increasing amount of smoking in pack-years showed a strong inverse linear trend with PD risk (p=0.005), as did number of cigarettes smoked per day (p=0.019). In general, the strengths of association were stronger and the confidence intervals generally narrowed when using our current adjudicated method compared to results when using the original algorithm.

In our secondary analyses in which we assessed the magnitude of residual confounding by time eligible in CMS, we observed estimates of association that were slightly attenuated to the null when we controlled for time eligible in CMS attenuated (online Table D). Results of the Cox regression analysis also showed similar associations for smoking status, pack-years, and cigarettes per day, although linear trends in the Cox models did not reach statistical significance (data not shown).

We assessed how our results would change if we had only CMS FFS to supplement CHS data, that is, with no neurologist adjudication using additional information. We found that 55.8% of PD identified by CMS FFS claims alone and 53.0% of PD identified by CHS data would be considered "probable PD" by movement disorders specialists. This proportion markedly increased to 72.7% when evidence of PD came from both CHS and CMS FFS. In associations with baseline smoking status, estimates were relatively strong for PD defined in CMS FFS alone, but were strongest when the definition of PD required evidence in both CHS and CMS (Table 3).

DISCUSSION

Using information from CMS claims and an adjudication process to enhance the identification of PD in the CHS allowed us to capture an additional 75 participants with PD who were not previously identified in the original algorithm[16] based on data available in CHS alone. However, the adjudication process also resulted in changing the PD status of 80 participants previously defined as having PD. Therefore, the baseline prevalence estimate and incidence rate for PD in the CHS were similar to what we previously reported.

The inverse association between smoking and PD is one of the most consistent findings in PD research. First observed in 1959,[21] this association has been replicated in meta-analyses and pooled analyses of more than 50 studies.[19, 22] Furthermore, prospective studies have replicated this observation,[23] providing evidence that selection and survival bias do not account for this inverse association. We have leveraged this well-established association in PD to gauge the utility of our case ascertainment methods. Although our enhanced PD

definition did not alter prevalence and incidence estimates, estimates of association with baseline smoking characteristics were strengthened compared to those based on the original algorithm, possibly as a result of reduced non-differential disease misclassification.[17] We also evaluated the utility of supplementing CHS data (self-report, medications, in-patient hospitalizations) with CMS outpatient claims without the benefit of medical records from hospitalization review, information on adjudicated deaths, or an adjudication process by expert reviewers. Defining PD as those with any evidence in both CHS <u>and</u> CMS claims uncovered the strongest estimates of association for baseline smoking characteristics relative to other PD definitions. These estimates approached those reported in a meta-analysis of cigarette smoking and PD across 48 studies.[19] In contrast, defining PD by CMS claims without evidence in CHS substantially obscured underlying associations with baseline smoking. Our results suggest that billing data are useful in identifying PD cases, but should be used with caution in the absence of other sources of data to establish the PD diagnosis with greater certainty. At the same time, our results also suggested that information in the CHS medical records of hospitalizations did not serve as an optimal source to obtain a gold standard definition for PD since the majority of the medical records lacked information to either confirm or refute a diagnosis of PD.

Several limitations deserve mention. The most significant limitation is the absence of a documented neurologic examination for PD by a neurologist or movement disorder specialist. Medical records from outpatient visits were not available. Although our age-specific baseline prevalence estimates fall within the range of estimates reported in other studies (typically between 0.2% to 1.2% for increasing age categories),[24, 25] our age-specific incidence estimates are within the higher range of what has been reported in the literature. A recent study in Kaiser Permanente Medical Care Program of Northern California reported a crude incidence rate of 190.5 per 100,000 person-years for their highest age category of 80-89 years [26] whereas a study in Northern Manhattan reported a rate of 213 per 100,000 person-years among subjects 80 years [27]. In CHS, we observed a rate of 273.2 per 100,000 person-years for 80-89 year olds, which is substantially greater, although not the highest rate reported in the literature. Several reviews of incidence rate of PD, which included European studies documented estimates for 80-89 age categories that ranged between 116 to 678 per 100,000 person-years.[28, 29] Likely, we over-diagnosed PD using our current methods, but without a clinical gold standard, the degree of misclassification remains unclear. Nonetheless, our enhanced case ascertainment method provided credible estimates of association with baseline smoking characteristics that were observed in other studies.[19, 22] Observing the expected inverse association of smoking with PD, in which estimates were similar to those derived from a large meta-analysis[19] provides greater assurance of the validity of our case ascertainment. Ultimately, however, without a clinical gold standard against which to compare our case ascertainment method, we are unable to ascertain the degree of validity conferred by this method. Such a resolution is possible if this case ascertainment method could be reproduced in a cohort that also contains a clinical gold standard such as the United Kingdom Parkinson's Disease Society Brain Bank Criteria.[30]

Our study is also limited by the inability to document services not captured in Medicare; these include certain services not covered by Medicare, and more importantly, services for those who are enrolled in managed care organizations since Medicare only provides claims for FFS.[31] A study assessing the external validity of the CHS documented that over 91.4% of CHS participants were enrolled in FFS,[32] suggesting that claims for most participants in the CHS would be captured in our study. Nonetheless, managed care increased over time and we would expect a decreasing proportion of CHS participants to remain in CMS FFS because of this trend. Second, the inverse association between smoking and PD observed in our study may be attributed, at least in part, by the healthcare utilization behavior of smokers. A study assessing the association between smoking behavior and use of primary

care services among 254,382 adults 45 years and older with universal healthcare in Australia found that current smokers were less likely than non-smokers to make a Medicare claim.[20] The direction of such a systematic behavior among smokers would bias the estimates downward, potentially artificially producing or at least strengthening inverse associations with current smoking. Such a bias would be most pronounced in studies that rely exclusively on Medicare claims for outcome ascertainment. Because we used other data sources to identify PD including self-report, annual medication inventory, and hospitalization discharge records, we do not expect our study to have been greatly influenced by this potential bias. Nonetheless, indirect validation by assessing whether an expected association is observed represents a suboptimal method due to the potential vulnerability to such biases.

Finally, although our enhanced method using out-patient billing codes allowed us to pinpoint a more accurate index date for PD than our previous algorithm, our incident date still appears to contain some degree of error. As expected, the new incident date using CMS claims occurred over four years earlier than the incident date defined by the previous algorithm. Nonetheless, the average age of onset with the new incident date was 79 years of age, at least ten years later than what should be expected for a date of onset defined by symptoms.[29] Clearly, our current incident date reflects patterns in healthcare utilization, or medication use rather than by clinical signs. Such misclassification affects our calculation of incidence rates as well as our Cox models, which provided similar estimates of association to logistic models but which were not statistically significant.

CONCLUSION

The recent linkage of the CMS data to the CHS has opened a number of research possibilities. In this study, we have used the CMS data to enhance our case ascertainment methods for PD, substantively contributing to screening, adjudication, and analysis. Although our enhanced PD ascertainment method did not importantly alter prevalence and incidence estimates, it may have reduced misclassification compared to our previous algorithm. Future PD research should take advantage of the CMS data in different ways, such as to provide a common framework to pool resources across other similar population-based cohorts and to address health utilization patterns and cost burden.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

REFERENCES

1. Ascherio A, Chen H, Schwarzschild MA, Zhang SM, Colditz GA, Speizer FE. Caffeine, postmenopausal estrogen, and risk of Parkinson's disease. Neurology. 2003; 60:790–5. [PubMed: 12629235]

2. Chen H, Schernhammer E, Schwarzschild MA, Ascherio A. A prospective study of night shift work, sleep duration, and risk of Parkinson's disease. Am J Epidemiol. 2006; 163:726–30. [PubMed: 16495472]

3. Chen H, Zhang SM, Hernan MA, Schwarzschild MA, Willett WC, Colditz GA, et al. Nonsteroidal anti-inflammatory drugs and the risk of Parkinson disease. Arch Neurol. 2003; 60:1059–64. [PubMed: 12925360]

4. Gao X, Chen H, Schwarzschild MA, Ascherio A. Use of ibuprofen and risk of Parkinson disease. Neurology. 2011; 76:863–9. [PubMed: 21368281]

5. Gao X, Simon KC, Schwarzschild MA, Ascherio A. Prospective study of statin use and risk of Parkinson disease. Arch Neurol. 69:380–4. [PubMed: 22410446]

6. Thacker EL, O'Reilly EJ, Weisskopf MG, Chen H, Schwarzschild MA, McCullough ML, et al. Temporal relationship between cigarette smoking and risk of Parkinson disease. Neurology. 2007; 68:764–8. [PubMed: 17339584]

7. Zhang SM, Hernan MA, Chen H, Spiegelman D, Willett WC, Ascherio A. Intakes of vitamins E and C, carotenoids, vitamin supplements, and PD risk. Neurology. 2002; 59:1161–9. [PubMed: 12391343]

8. Simon KC, Chen H, Schwarzschild M, Ascherio A. Hypertension, hypercholesterolemia, diabetes, and risk of Parkinson disease. Neurology. 2007; 69:1688–95. [PubMed: 17761552]

9. Hernan MA, Chen H, Schwarzschild MA, Ascherio A. Alcohol consumption and the incidence of Parkinson's disease. Ann Neurol. 2003; 54:170–5. [PubMed: 12891669]

10. Palacios N, Gao X, McCullough ML, Jacobs EJ, Patel AV, Mayo T, et al. Obesity, diabetes, and risk of Parkinson's disease. Mov Disord. 2011; 26:2253–9. [PubMed: 21739472]

11. Palacios N, Gao X, O'Reilly E, Schwarzschild M, McCullough ML, Mayo T, et al. Alcohol and risk of Parkinson's disease in a large, prospective cohort of men and women. Mov Disord. 27:980–7. [PubMed: 22714720]

12. Dawber TR, Meadors GF, Moore FE Jr. Epidemiological approaches to heart disease: the Framingham Study. Am J Public Health Nations Health. 1951; 41:279–81. [PubMed: 14819398]

13. ARIC Investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. 1989; 129:687–702.

14. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, et al. Multi-ethnic study of atherosclerosis: objectives and design. Am J Epidemiol. 2002; 156:871–81. [PubMed: 12397006]

15. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, et al. The Cardiovascular Health Study: design and rationale. Ann Epidemiol. 1991; 1:263–76. [PubMed: 1669507]

16. Ton TG, Jain S, Boudreau R, Thacker EL, Strotmeyer ES, Newman AB, et al. Post hoc Parkinson's disease: identifying an uncommon disease in the Cardiovascular Health Study. Neuroepidemiology. 2010; 35:241–9. [PubMed: 20881426]

17. Rothman, KJ.; Greenland, S.; Lash, TL. 3rd. Wolters Kluwer Health/Lippincott Williams & Wilkins; Philadelphia: 2008. Modern epidemiology.

18. Medicare: A Primer. The Henry J. Kaiser Family Foundation; Menlo Park: 2010. Contract No.: Document Number|

19. 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; 52:276–84. [PubMed: 12205639]

20. Jorm LR, Shepherd LC, Rogers KD, Blyth FM. Smoking and use of primary care services: findings from a population-based cohort study linked with administrative claims data. BMC health services research. 2012; 12:263. [PubMed: 22900643]

21. Dorn HF. Tobacco consumption and mortality from cancer and other diseases. Public Health Rep. 1959; 74:581–93. [PubMed: 13668008]

22. Ritz B, Ascherio A, Checkoway H, Marder KS, Nelson LM, Rocca WA, et al. Pooled analysis of tobacco use and risk of Parkinson disease. Arch Neurol. 2007; 64:990–7. [PubMed: 17620489]

23. Chen H, Huang X, Guo X, Mailman RB, Park Y, Kamel F, et al. Smoking duration, intensity, and risk of Parkinson disease. Neurology. 74:878–84. [PubMed: 20220126]

24. Fall PA, Axelson O, Fredriksson M, Hansson G, Lindvall B, Olsson JE, et al. Age-standardized incidence and prevalence of Parkinson's disease in a Swedish community. J Clin Epidemiol. 1996; 49:637–41. [PubMed: 8656224]

25. von Campenhausen S, Bornschein B, Wick R, Botzel K, Sampaio C, Poewe W, et al. Prevalence and incidence of Parkinson's disease in Europe. Eur Neuropsychopharmacol. 2005; 15:473–90. [PubMed: 15963700]

26. Van Den Eeden SK, Tanner CM, Bernstein AL, Fross RD, Leimpeter A, Bloch DA, et al. Incidence of Parkinson's disease: variation by age, gender, and race/ethnicity. Am J Epidemiol. 2003; 157:1015–22. [PubMed: 12777365]

27. Mayeux R, Marder K, Cote LJ, Denaro J, Hemenegildo N, Mejia H, et al. The frequency of idiopathic Parkinson's disease by age, ethnic group, and sex in northern Manhattan, 1988-1993. Am J Epidemiol. 1995; 142:820–7. [PubMed: 7572958]

28. von Campenhausen S, Bornschein B, Wick R, Botzel K, Sampaio C, Poewe W, et al. Prevalence and incidence of Parkinson's disease in Europe. European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology. 2005; 15:473–90. [PubMed: 15963700]

29. Twelves D, Perkins KS, Counsell C. Systematic review of incidence studies of Parkinson's disease. Mov Disord. 2003; 18:19–31. [PubMed: 12518297]

30. National Collaborating Centre for Chronic Conditions (UK). Parkinson's Disease: National Clinical Guideline for Diagnosis and Management in Primary and Secondary Care. Royal College of Physicians (UK); London: 2006. (NICE Clinical Guidelines, No. 25.) 5, Diagnosing Parkinson's disease. Available from: http://www.ncbi.nlm.nih.gov/books/NBK48502/

31. Warren JL, Klabunde CN, Schrag D, Bach PB, Riley GF. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. Medical care. 2002; 40:IV–3. 18.

32. DiMartino LD, Hammill BG, Curtis LH, Gottdiener JS, Manolio TA, Powe NR, et al. External validity of the cardiovascular health study: a comparison with the Medicare population. Medical care. 2009; 47:916–23. [PubMed: 19597373]

Key points

- Supplementing information using Medicare claims data to identify Parkinson's disease contributed to a decrease in disease misclassification in an on-going cohort study of cardiovascular disease

- Indirect validation showed strongest associations with baseline smoking characteristics when Parkinson's disease was defined by evidence that came from both study data AND Medicare claims.

- Use of Medicare claims data in on-going cohort studies such as the Cardiovascular Health Study can open research possibilities for conditions such as Parkinson's disease for which large population-based longitudinal cohort studies are not feasible to establish.
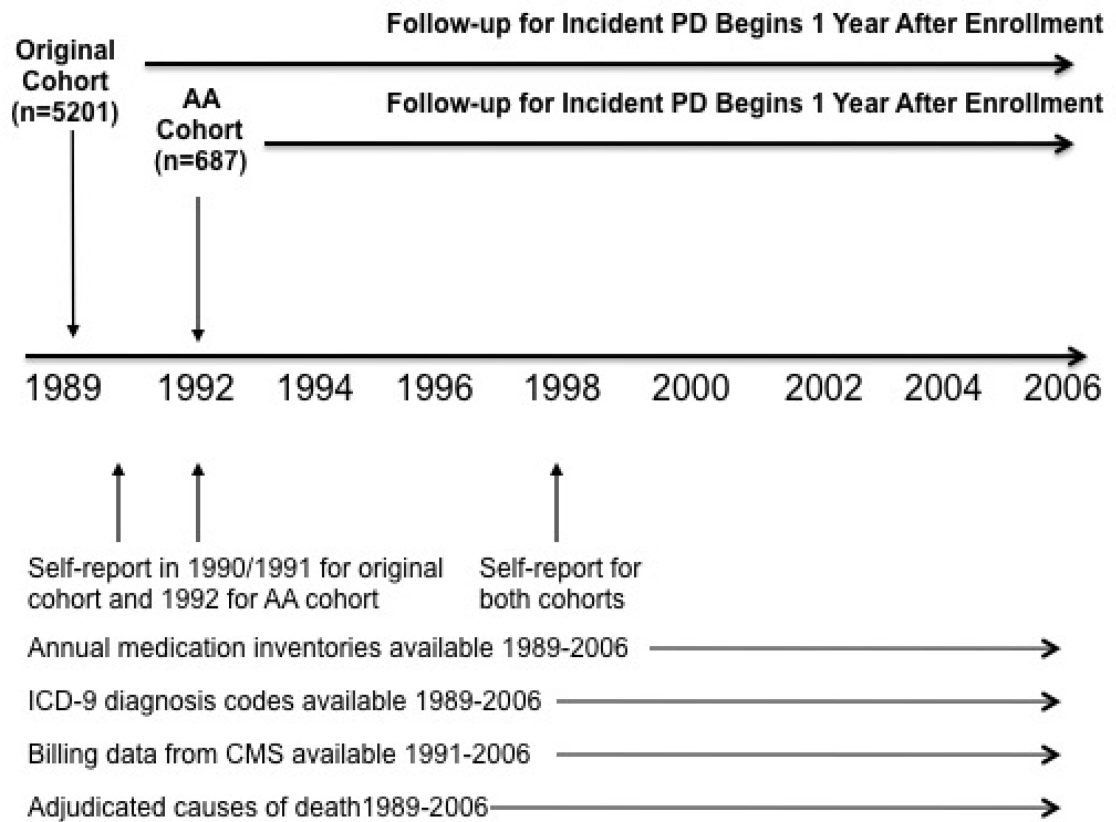
Figure 1.
Data sources for Parkinson's Disease Follow-up
AA = African American; CMS = Centers for Medicare and Medicaid Services; ICD-9 = International
Classification of Diseases – Ninth Version; PD = Parkinson's disease

Table 1

Incidence rate of PD over 70,796 person-years of follow-up years In the Cardiovascular Health Study, 1989-2007.

Characteristic	n	Person-years	Incidence Rate Per 100,000 person-years	95% CI	P-value**
Total	165	70,796	233.1	(200.1, 271.5)	
Age					0.328
65-69	61	26,709	228.4	(177.7, 293.5)	
70-79	89	32,996	269.7	(219.1,332.0)	
80+	15	5,293	283.4	(170.8, 470.0)	
Sex					<0.001
Men	95	26,853	353.8	(289.3, 432.6)	
Women	70	43,943	159.3	(126.0, 201.3)	
Education					0.9
<HS	43	18,794	228.8	(169.7, 308.5)	
HS	50	20,261	246.8	(187.0, 325.6)	
>HS	72	31,519	228.4	(181.3, 287.8)	
Race					0.8
White/Asian/AIANs	143	60,703	235.6	(200.0, 277.5)	
African American	22	10,094	217.9	(143.5, 331.0)	

* trend test for age; Pearson chi-square test for sex and race

**
  log rank test

Table 2

Comparison of Associations from Multivariate Logistic Models of incident PD[*] and Baseline Smoking Characteristics Using Current Case Ascertainment Method and Original Algorithm in the Cardiovascular Health Study, 1989-2006.

| Characteristic | Crude Distribution | | | | Current Adjudication | | | Original Algorithm[†] | | |
	PD n=165	(%)	no PD n=5679	(%)	OR[**]	Adjusted[**] 95% CI	P-value[a]	OR[‡]	Adjusted[‡] 95% CI	P-value[a]
Smoking Status							0.009			0.025
Never	80	(48.5)	2634	(46.4)	1.00	Reference		1.00	Reference	
Former	74	(44.9)	2297	(40.5)	0.83	(0.59, 1.16)		0.86	(0.61, 1.12)	
Current	11	(6.7)	742	(13.1)	0.42	(0.22, 0.79)		0.46	(0.24, 0.88)	
Years since quitting							0.005			0.015
Never smoker	80	(49.4)	2634	(47.0)	1.00	Reference		1.00	Reference	
30 years	23	(14.2)	604	(10.8)	0.99	(0.61, 1.61)		1.01	(0.61, 1.68)	
20-29	21	(13.0)	560	(10.0)	0.96	(0.58, 1.58)		0.92	(0.54, 1.58)	
10-19	12	(7.4)	585	(10.4)	0.54	(0.29, 1.00)		0.76	(0.43, 1.34)	
>1-10	15	(9.2)	477	(8.5)	0.84	(0.48, 1.48)		0.40	(0.43, 1.40)	
Current smoker	11	(6.8)	742	(13.3)	0.42	(0.22, 0.79)		0.46	(0.24, 0.88)	
Cigarettes per day							0.019			0.041
Never smoker	80	(50.3)	2634	(47.5)	1.00	Reference		1.00	Reference	
<10	16	(10.1)	624	(11.2)	0.80	(0.46, 1.38)		0.90	(0.52, 1.54)	
10-19	23	(14.5)	802	(14.4)	0.81	(0.51, 1.31)		0.94	(0.59, 1.51)	
20-28	29	(18.2)	935	(16.5)	0.76	(0.48, 1.19)		0.62	(0.37, 1.03)	
29	11	(6.9)	559	(10.0)	0.45	(0.51, 1.31)		0.61	(0.33, 1.14)	
Total pack-years							0.005			0.032
Never smoker	80	(50.6)	2634	(48.0)	1.00	Reference		1.00	Reference	
1st quartile(0-13)	29	(18.4)	829	(15.1)	0.99	(0.64, 1.54)		0.82	(0.57, 1.50)	
2nd quartile (14-27)	17	(10.8)	643	(11.7)	0.72	(0.42, 1.24)		0.95	(0.58, 1.56)	
3rd quartile (28-49)	13	(8.3)	505	(9.2)	0.65	(0.36, 1.20)		0.72	(0.42, 1.23)	
4th quartile (>50)	19	(12.0)	880	(16.0)	0.51	(0.30, 0.85)		0.53	(0.29, 0.96)	

*
PD defined as Probable PD by adjudication

**
Adjusted for age, sex, African American race, and education

[†]
Incident PD = 154, as described in reference 2;

[‡]
Adjusted for age, sex, African American race, and education.

[a]
Test for linear trend

Table 3

Comparison of Associations* with Baseline Smoking Status by PD** Defined by Various Data Sources, Restricted to Time Period of Original Algorithm, 1989-2001.

Characteristic	PD in CMS (n=267)			PD in CHS (n=254)			PD in either CMS or CHS (n=367)			PD in Both CMS or CHS (n=254)		
	OR	95% CI	P-value	OR	95% CI	P-value	OR	95% CI	P-value	OR	95% CI	P-value
Smoking Status			0.199			0.052			0.125			0.066
Never	1.00	Reference		1.00	Reference		1.00	Reference		1.00	Reference	
Former	0.86	(0.66, 1.14)		0.89	(0.67, 1.18)		0.92	(0.73, 1.16)		0.78	(0.56, 1.34)	
Current	0.79	(0.52, 1.21)		0.62	(0.38, 0.98)		0.74	(0.51, 1.08)		0.61	(0.33, 1.11)	
Years since quitting			0.285			0.021			0.064			0.037
Never smoker	1.00	Reference		1.00	Reference		1.00	Reference		1.00	Reference	
30 years	0.97	(0.65, 1.44)		1.13	(0.77, 1.68)		1.09	(0.78, 1.53)		0.96	(0.58, 1.58)	
20-29	1.04	(0.70, 1.57)		0.81	(0.52, 1.27)		0.95	(0.66, 1.37)		0.88	(0.52, 1.50)	
10-19	0.66	(0.41, 1.06)		0.88	(0.57, 1.36)		0.80	(0.54, 1.17)		0.70	(0.39, 1.26)	
>1-10	0.75	(0.53, 1.37)		0.74	(0.44, 1.23)		0.86	(0.57, 1.30)		0.68	(0.33, 1.27)	
Current smoker	0.79	(0.52, 1.21)		0.62	(0.28, 0.98)		0.74	(0.51, 1.08)		0.61	(0.33, 1.11)	
Total pack-years			0.052			0.022			0.058			0.010
Never smoker	1.00	Reference		1.00	Reference		1.00	Reference		1.00	Reference	
1st quartile(0-13)	1.05	(0.73, 1.50)		0.96	(0.66, 1.39)		1.03	(0.76, 1.40)		0.95	(0.60, 1.52)	
2nd quartile (14-27)	0.96	(0.64, 1.43)		1.08	(0.73, 1.60)		1.00	(0.71, 1.41)		1.06	(0.65, 1.74)	
3rd quartile (28-49)	0.93	(0.60, 1.46)		1.01	(0.65, 1.56)		1.03	(0.71, 1.49)		0.84	(0.47, 1.49)	
4th quartile (50)	0.64	(0.43, 0.96)		0.52	(0.33, 0.80)		0.66	(0.46, 0.93)		0.42	(0.23, 0.76)	
Cigarettes per day			0.247			0.061			0.238			0.036
Never smoker	1.00	Reference		1.00	Reference		1.00	Reference		1.00	Reference	
<10	1.00	(0.66, 1.52)		0.85	(0.55, 1.33)		0.88	(0.61, 1.28)		1.02	(0.60, 1.73)	
10-19	0.85	(0.58, 1.25)		1.06	(0.74, 1.53)		1.05	(0.77, 1.43)		0.75	(0.45, 1.25)	
20-28	0.88	(0.62, 1.26)		0.82	(0.57, 1.19)		0.86	(0.63, 1.17)		0.84	(0.54, 1.32)	
29	0.80	(0.51, 1.25)		0.55	(0.33, 0.93)		0.78	(0.53, 1.16)		0.46	(0.23, 0.91)	

* Adjusted for age, sex, African American race, and education

*
** includes both prevalent and incident PD