

Estimation of High-Dimensional Propensity Scores with Multiple Exposure Levels

Short Title: Multilevel High-Dimensional Propensity Scores

Authors: Maria Eberg MSc^{1*}, Robert W. Platt PhD^{1,2,3,4}, Pauline Reynier MSc¹, Kristian B. Filion PhD^{1,3,5},

Author affiliations:

¹ Centre for Clinical Epidemiology, Lady Davis Research Institute, Jewish General Hospital, Montreal, Quebec, Canada

² The Research Institute of the McGill University Health Centre, Montreal, Quebec, Canada.

³ Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Quebec, Canada

⁴ Department of Pediatrics, McGill University, Montreal, Quebec, Canada

⁵ Department of Medicine, McGill University, Montreal, Quebec, Canada

* Maria Eberg is currently an employee at Cancer Care Ontario, Toronto, Ontario. She conducted this work while an employee at the Centre for Clinical Epidemiology of the Lady Davis Research Institute.

Word Count: 2,414; **Tables:** 2; **Figures:** 1;

Address for Correspondence:

Kristian B. Filion, PhD

Associate Professor and William Dawson Scholar

Departments of Medicine and of Epidemiology, Biostatistics, and Occupational Health, McGill University

Centre For Clinical Epidemiology, Lady Davis Research Institute - Jewish General Hospital

3755 Cote Ste-Catherine, H-410.1

Montreal, Quebec H3T 1E2 Canada

Tel: (514) 340-8222 x 28394

Fax: (514) 340-7564

E-mail: kristian.filion@mcgill.ca

Keywords: high-dimensional propensity scores; comparative effectiveness; pharmacoepidemiology; multilevel exposure

Key points:

- In settings with >2 exposure levels, multinomial high-dimensional propensity scores produce comparable results to traditional binomial propensity scores applied to all pairwise comparisons.

- Multinomial high-dimensional propensity scores may be preferable when investigators want to simultaneously estimate propensity scores for all treatment levels in a single cohort.

Research sponsors: This study was conducted using data acquired thanks to an operating grant from the Canadian Institutes of Health Research (grant number MOP 119433) and with the support of the Canadian Network for Observational Drug Effect Studies (CNODES), a collaborating center of the Drug Safety and Effectiveness Network (DSEN) that is funded by the CIHR (Grant Number DSE-146021). Dr. Filion is supported by a salary support award from the *Fonds de recherche du Québec – santé* (FRQS; Quebec Foundation for Health Research) and a William Dawson Scholar award from McGill University.

Conflicts of interest: RWP has received personal fees from Amgen, Abbvie, Pfizer, Eli Lilly, and Novartis. Ms. Eberg, Ms. Reynier, and Dr. Filion declare no competing interests.

Statement about prior posting: This provided manuscript is the sole work of the identified contributing authors. This methodological study is an extension of substantive work that has been previously presented as an abstract* and is currently being considered for publication. This methodological work has never been submitted to academic journals or shared with a public audience.

*Filion KB, Dell'Aniello S, Eberg M, Renoux C, Daskalopoulou SS, Suissa S. Varenicline and the risk of adverse cardiovascular events: a population-based cohort study. *Circulation* 2014; 129:AP135. 2014 American Heart Association's Joint – 54th Cardiovascular Disease Epidemiology and Prevention - and - Nutrition, Physical Activity and Metabolism Conference. March 18-22, 2014. San Francisco, California.

ABSTRACT

Purpose: Little information is available on the performance of high-dimensional propensity scores (HDPS) in settings with >2 exposure levels. Our objective was to adapt the HDPS algorithm to allow for the inclusion of multilevel treatments and compare estimates obtained via this approach to those obtained via pairwise comparisons in a case study using real-world data.

Methods: We conducted a retrospective cohort study of cardiovascular events associated with three smoking cessation drugs (varenicline, bupropion, nicotine replacement therapy (NRT)) using the Clinical Practice Research Datalink. We applied the binary HDPS algorithm adjusted for pre-specified and empirically-selected covariates to cohorts formed by each treatment pair. We then constructed multinomial HDPS models on a cohort of incident users of any of the three drugs, adjusting for pre-defined covariates and different combinations of empirically-selected covariates. After trimming the area of non-overlap of the HDPS distributions, the effects of the study drugs on cardiovascular events were estimated with Cox proportional hazards models adjusted for propensity score category.

Results: Outcome models adjusted for multinomial HDPS estimated treatment effects that were slightly more protective than those estimated in pairwise comparisons (varenicline vs NRT: $HR_{Multinomial} = 0.60-0.62$, $HR_{Pairwise} = 0.64$; bupropion vs NRT: $HR_{Multinomial} = 0.70-0.72$, $HR_{Pairwise} = 0.76$). Trimming rates were similar between the two approaches.

Conclusions: The extension of HDPS to multilevel exposures is a valid and practical approach to confounder control that may be useful when comparing different classes of drugs prescribed for the same indication or different molecules within a given drug class.

INTRODUCTION

The high-dimensional propensity score (HDPS) algorithm uses an automated, data-driven approach to augment traditional propensity score model with a set of empirically-derived confounders.¹ The goal of using the algorithm is to improve the estimation of a patient's probability to receive one treatment versus an alternative treatment for the purposes of adjustment, matching, or weighting in the outcome model. Schneeweiss et al. have described the main steps for defining and implementing the HDPS using administrative data.¹ Briefly, the algorithm first scans predetermined dimensions of administrative data, and for each dimension, it selects the n most common codes. It then assesses the recurrence of each identified code and creates three covariates based on the frequency with which a given code occurs for every individual in the study cohort. The algorithm ranks these empirically-derived covariates across dimensions according to their potential for confounding control. For each candidate variable, a 'bias-inflation factor' is calculated as a function of the imbalance in prevalence of the candidate confounder among exposed and unexposed subjects as well as the independent association between the candidate variable and the study outcome. Finally, the top k covariates are selected for inclusion in the propensity score model along with variables prespecified by the investigator.

The HDPS algorithm is not designed for multi-category exposures that may be encountered in studies of comparative effectiveness or safety, and little information exists on the performance of HDPS algorithm in these settings. When encountering >2 exposure levels, investigators typically estimate HDPS for all pairwise comparisons.^{2,3,4} That is, in a situation where a subject may be exposed to one of three drug categories, A, B, or C, the conventional use of the algorithm would result in three pairwise comparisons: A vs B, B vs C, and A vs C. In such cases, implementation of the HDPS algorithm results in models with three different sets of covariates empirically selected

by the algorithm. As a result, trimming of the initial population to groups with overlapping propensity scores may result in different populations for each comparison. The inclusion of different target populations could cause nontransitive estimates of the exposure effect (e.g., A is better than B, B is better than C, but A is worse than C).⁵

If the comparisons of interest are targeted against one reference group, the pairwise approach may be acceptable since only the pairwise contrasts against the reference group are relevant. However, an investigator may be interested in comparing several treatments among subjects who had some chance of receiving any one of several treatment options.⁶ For instance, Spreeuwenberg et al. simultaneously evaluated five psychotherapy interventions. In this setting, the probability of each treatment assignment can be estimated using a model for a categorical outcome such as multinomial logistic regression.⁷ The advantage of using one model to calculate treatment propensities versus separate models for each contrast is the simultaneous estimation of multiple exposures. The creation of a single study cohort that results from the use of a multinomial propensity model is uniquely consistent with “empirical equipoise”, the condition that patients should be plausible candidates for each treatment under study.⁸

Our objective was to adapt the HDPS algorithm to allow for the inclusion of multilevel treatments and to compare estimates obtained via this approach to those obtained via pairwise comparisons of each contrast. We applied both HDPS methods to a case study of smoking cessation drugs and the risk of cardiovascular events using real-world data. We used several sets of empirically-selected covariates in multinomial HDPS models to explore different methods of variable selection with three exposure levels.

METHODS

Data Source

We conducted a population-based cohort study using data from the Clinical Practice Research Datalink (CPRD). The CPRD, which has been described in detail elsewhere⁹, is a population-based clinical database that contains detailed clinical records for more than 15 million patients seen at over 675 general practitioner practices in the UK. Approximately 58% of CPRD data are linkable to the Hospital Episode Statistics (HES) database¹⁰, which contains detailed hospitalization records. This study was approved by the Independent Scientific Advisory Committee of the CPRD (protocol number 13_045RA2 v2) and the Research Ethics Board of the Jewish General Hospital in Montreal, Canada.

Case Study: Smoking Cessation Drugs and the Risk of Adverse Cardiovascular Events

The case study was designed to assess the association between three smoking cessation drugs and the risk of adverse cardiovascular events. We formed a cohort of new users of smoking cessation therapy ≥ 18 years old who received a new prescription for bupropion, varenicline, or nicotine replacement therapy (NRT) between December 1st, 2006 and August 31th, 2012. We excluded patients who had: < 1 year of medical history, a prescription for one of the smoking cessation drugs in the year before cohort entry, and those with a cardiovascular event in the year before cohort entry. The primary endpoint was a composite of myocardial infarction, coronary revascularization, stroke, and all-cause mortality, recorded in CPRD, with CPRD data supplemented by HES for patients for whom linkage was available. Patients were followed from their new prescription for a smoking cessation drug until the occurrence of an event or censoring due to discontinuation of the study medication, departure from the CPRD, end of follow-up (1 year), or end of the study period (December 31, 2012), whichever occurred first. The data dimensions for the HDPS algorithm included diagnoses, procedures, prescriptions, demographic characteristics, and general

practitioner practice characteristics, with covariates assessed in the year before cohort entry. An “as-treated” exposure definition was used, with patients censored 7 days after the end of their last prescription or upon switching drugs.

Statistical Analyses

We applied the HDPS algorithm to each pairwise contrast. The k top ranked covariates were selected empirically for each of the comparisons. To avoid overfitting, k was chosen according to the rule $m/20 < k < m/10$, where m is the sample size of the smallest exposure group, resulting in $k=500$.¹¹ For pairwise contrasts, our cohorts were comprised of only patients receiving one of the two relevant drugs. Pairwise HDPS models were adjusted for pre-defined variables and covariates empirically selected by the algorithm for the respective comparison. In the multinomial analyses, we used the whole study population of new users of one of the three smoking cessation drugs and adjusted HDPS models for pre-defined covariates and different combinations of k empirically selected covariates. Predefined confounders were measured at or before cohort entry and included: age, sex, geographic region, calendar year of cohort entry, lifestyle variables (alcohol use, obesity), comorbidities (atrial fibrillation, acute coronary syndrome and/or coronary artery disease, cerebrovascular disease, congestive heart failure, chronic obstructive pulmonary disorder, diabetes, depression, hypertension, hypercholesterolemia, peripheral vascular disease, previous myocardial infarction, previous coronary revascularization, previous stroke), previously prescribed drugs (aspirin, angiotensin-converting enzyme inhibitors, angiotensin receptor blockers, beta-blockers, calcium-channel blockers, diuretics, fibrates, non-steroidal anti-inflammatory drugs, statins), number of drug classes prescribed in the previous year, and an indicator variable for whether the patient was linkable to HES.

Ranking and selection of the empirically identified covariates was performed separately for each pairwise comparison, and all duplicate variables were eliminated. In multinomial HDPS models, the combinations of the empirically selected covariates included: 1) covariates selected by all pairwise comparisons (concordant covariates); 2) top k covariates, including all concordant covariates and the next top ranked covariates from each comparison; and 3) all the top ranked covariates from each contrast. The multinomial logistic models estimated two propensity scores for each patient: one for each medication versus the referent category. For both multinomial and pairwise analyses, the area of non-overlap was trimmed from the propensity score distributions. In the case of pairwise contrasts, we adjusted the outcome models for age, sex, and propensity score deciles; in case of the multinomial treatment models, we adjusted for age, sex, and propensity score quintiles (eight dummy variables; four for each of the two propensity scores). To evaluate the degree of confounding and additional adjustment attributable to HDPS, we compared the results from the outcome models using HDPS for adjustment to the estimates from the unadjusted outcome models and models with propensity scores estimated with pre-defined covariates only.

RESULTS

Our base study population included 233,738 patients receiving a new prescription for a smoking cessation drug: 68,468 varenicline users, 9,931 bupropion users, and 155,339 NRT users (Figure 1). A total of 90 events occurred among varenicline users, 13 events among bupropion users, and 513 events among NRT users. The distribution of baseline characteristics was similar among varenicline and bupropion users (Table 1). In comparison, NRT users were substantially older than members of the other treatment groups and had a higher prevalence of comorbidities and medication use.

The difference in the HDPS-adjusted hazard ratios, as compared to the crude hazard ratios, ranged from 0.01 to 0.37 for the pairwise comparisons and from 0.17 to 0.33 for the multinomial comparisons. The inclusion of empirically-selected variables in the propensity score model resulted in an additional change in the estimated hazard ratios of 0.05 – 0.08 as compared to adjustment with pre-defined covariates only. Overall, estimated treatment effects were similar for each analytic strategy (Table 2). Treatment effects of bupropion and varenicline estimated by multinomial HDPS were slightly more protective as compared to the results of pairwise comparisons. However, the differences were not clinically important. The use of multinomial models to estimate HDPS resulted in a higher proportion of subjects excluded from the outcome model for all analyses except for the multinomial HDPS analysis including only the concordant empirically-selected variables (Table 2, Analysis 1). As mentioned above, the NRT population was substantially different from the comparator groups, and across all the HDPS comparisons, there was a higher proportion of NRT users trimmed due to non-overlap of the HDPS distribution. Hazard ratios for varenicline were similar between the pairwise contrasts and the multinomial propensity score analyses. The results for bupropion were more heterogeneous, likely due to the

small number of bupropion-exposed events; there was little statistical power to further investigate the observed differences.

DISCUSSION

HDPS represents one of the most popular methods to reduce confounding and selection bias in pharmacoepidemiologic research, though its application has historically been limited to two treatment categories. The current study builds on previous research on generalized propensity scores and extends the evaluation of HDPS performance to a multilevel treatment setting.^{5,12,13,14,15} We applied HDPS to a setting with three exposures and compared estimates obtained via multinomial HDPS models adjusted for combinations of empirically-selected covariates to those obtained via pairwise comparisons of each possible contrast. In this case study, we observed no important differences between the single HDPS models (multinomial logistic models) versus separate HDPS models (separate binomial logistic models).

Similar findings have been documented in a previous study of general propensity score estimation and adjustment.¹³ In our study, slightly more subjects were excluded due to non-overlap of the HDPS distributions when using multinomial HDPS models. However, it may be expected that in any three-way comparison, the area of three-way “equipoise” will be smaller than the area for any two-way comparisons. The pairwise comparison method was as valid as the more complex multinomial HDPS algorithm, and it may generally be preferred due to its straightforward application. However, it does not allow for direct comparison of all treatment categories. The multinomial approach involves the estimation of fewer treatment and outcome models, and it uses only one population to compare the treatments. Multinomial modeling ensures appropriate estimation of treatment probabilities and provides higher efficiency in parameter estimation as compared to pairwise model fitting.⁸ For each participant, one can estimate a vector of all treatment probabilities at once and compare to individuals with similar generalized propensity scores $(\{\widehat{p(A)}, \widehat{p(B)}, \widehat{p(C)}\}_i = \{\widehat{p(A)}, \widehat{p(B)}, \widehat{p(C)}\}_j$ for individuals i and j). In contrast, pairwise

comparisons may include subjects with similar probability of receiving drug A, but disregard the differences in probability of receiving other drugs ($\{\widehat{p(A)}, \widehat{p(B)}, \widehat{p(C)}\}_j \neq \{\widehat{p(A)}, \widehat{p(B)}, \widehat{p(C)}\}_k$ while $\widehat{p(A)}_j = \widehat{p(A)}_k$ for individuals j and k).¹⁶ Moreover, it may be easier to test that the propensity score assumptions are satisfied using the estimates from a single model. Regardless of which modeling approach is used for propensity score estimation, covariate balance may be assessed through the use of quintile regression and standardized differences (if matching or weighting by HDPS) or weighted conditional standardized differences (if adjusting for HDPS)^{17,18}. The multinomial approach may be particularly useful when comparing different classes of drugs prescribed for the same indication or within class comparisons of different molecules.^{19,20} For instance, Leonard et al. used a variation of HDPS to estimate the associations between concomitant use of warfarin plus one of eight anti-hyperlipidemic drugs and the risk of gastrointestinal bleeding and intracranial hemorrhage.¹⁵

We should note that the use of one model to estimate propensity scores for >2 agents implies that individuals under study are plausible candidates for each study drug. However, the assumption of empirical equipoise for multiple agents may be inappropriate in some situations, for instance, in comparisons of monotherapy vs combination therapy, or in evaluations where the potency of the molecules within a class differs, such as comparison of atorvastatin against other, lower potency statins.

Our study has some potential limitations. We restricted our multinomial analyses to three combinations of empirically-selected covariates, though other combinations are possible. However, the results did not vary much between the multinomial analyses performed, and it is reasonable to assume that similar estimates would have been obtained using other permutations of

empirically-selected covariates. Also, we did not explore other methods of confounding control such as matching or weighting on propensity scores. In addition, this case study involved a three-way comparison; the discrepancy between binomial and multinomial HDPS may grow as the number of treatment levels increases. We adjusted for propensity score decile in our pairwise comparisons and propensity score quintiles for the multinomial analyses; the former may result in better confounding control. However, due to low event counts among bupropion and varenicline users and the potential for overfitting, we preferred to keep a similar number of parameters used in the outcome models. Finally, our use of an as-treated exposure definition resulted in a short duration of follow-up given the short prescription of smoking cessation drugs. Consequently, our study included a modest number of events, resulting in imprecise effect estimates.

CONCLUSIONS

We have demonstrated a possible implementation of a multinomial HDPS algorithm. We have presented a novel HDPS method that compares three treatment choices simultaneously and contrasted our method with the standard approach. The results obtained from various multinomial analyses and pairwise treatment comparisons were similar. We believe that the extension of the HDPS algorithm to multiple exposure categories could be practical when one wants to simultaneously compare multiple treatment categories.

DISCLOSURES

RWP has received personal fees from Amgen, Abbvie, Pfizer, and Novartis. Ms. Eberg, Ms. Reynier, and Dr. Filion declare no competing interests.

ACKNOWLEDGEMENTS

This study was conducted using data acquired thanks to an operating grant from the Canadian Institutes of Health Research (grant number MOP 119433) and with the support of the Canadian Network for Observational Drug Effect Studies (CNODES), a collaborating center of the Drug Safety and Effectiveness Network (DSEN) that is funded by the CIHR (Grant Number DSE-146021). Dr. Filion is supported by a salary support award from the Fonds de recherche du Québec – santé (FRQS; Quebec Foundation for Health Research) and a William Dawson Scholar award from McGill University. Robert Platt holds the Albert Boehringer I Chair in Pharmacoepidemiology at McGill University. The authors thank Mr. Matthew H Secrest for his help in editing the manuscript.

REFERENCES

1. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiol* 2009;20(4):512–522.
2. Zakai NA, Walker RF, MacLehose RF, Adam TJ, Alonso A, Lutsey PL. Impact of anticoagulant choice on hospitalized bleeding risk when treating cancer-associated venous thromboembolism. *J Thromb Haemost* 2018; 16:2403–2412.
3. Kollhorst B, Jobski K, Krappweis J, Schink T, Garbe E, Schmedt N. Antidepressants and the risk of death in older patients with depression: A population-based cohort study. *PLoS One*. 2019;14(4):e0215289. Published 2019 Apr 15. doi:10.1371/journal.pone.0215289.
4. Norby FL, Bengtson LGS, Lutsey PL, Chen LY, MacLehose RF, Chamberlain AM, et al. Comparative effectiveness of rivaroxaban versus warfarin or dabigatran for the treatment of patients with non-valvular atrial fibrillation. *BMC Cardiovascular Disord*. 2017; 17:238
5. Rassen JA, Shelat AA, Franklin JM, et al. Matching by propensity score in cohort studies with three treatment groups. *Epidemiol* 2013;24(3):401–409.
6. Yoshida K, Solomon DH, Haneuse S, Kim SC, Patorno E, Tedeschi SK, et al. Multinomial extension of propensity score trimming methods: a simulation study. *Am J Epidemiol*, 2019;188(3):609–616
7. Spreeuwenberg M, Bartak A, Croon M, et al. The multiple propensity score as control for bias in the comparison of more than two treatment arms. *Med Care*. 2010; 48:166-174.

8. Walker AM, Patrick AR, Lauer MS, et al. A tool for assessing the feasibility of comparative effectiveness research. *Comp Eff Res* 2013;3:11-20.
9. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data resource profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 2015;44:827-36.
10. Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D, Hardelid P. Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (HES APC). *Int J Epidemiol* 2017;46:1093-1093i.
11. Harrell FE. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. 1st ed. New York: Springer; 2001.
12. Imbens GW. The role of the propensity score in estimating dose-response functions. *Biometrika* 2000;87:706-10.
13. Cadarette SM, Gagne JJ, Solomon DH, Katz JN, Stürmer T. Confounder summary scores when comparing the effects of multiple drug exposures. *Pharmacoepidemiol Drug Saf* 2010;19:2-9.
14. McCaffrey DF, Griffin BA, Almirall D, Slaughter ME, Ramchand R, Burgette LF. A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Stat Med*. 2013;32(19):3388–3414. doi:10.1002/sim.5753
15. Leonard CE, Brensinger CM, Bilker WB, Kimmel SE, Han X, Nam YH, Gagne JJ, Mangaali MJ, Hennessy S. Gastrointestinal bleeding and intracranial hemorrhage in concomitant users of warfarin and antihyperlipidemics. *Int J Cardiol*. 2017; 228:761–770.

16. Lopez M, Gutman R. Estimation of causal effects with multiple treatments: a review and new ideas. *Stat Sci* 2017;32(3): 432-454
17. Leonard CE, Brensinger CM, Aquilante CL, Bilker WB, Boudreau DM, Deo R, et al. Comparative safety of sulfonylureas and the risk of sudden cardiac arrest and ventricular arrhythmia. *Diabetes Care*. 2018;41(4):713-722.
18. Austin PC. Goodness-of-fit diagnostics for the propensity score model when estimating treatment effects using covariate adjustment with the propensity score. *Pharmacoepidemiol Drug Saf*. 2008;17(12):1202-17
19. Coupland C, Dhiman P, Morriss R, Arthur A, Barton G, Hippisley-Cox J et al. Antidepressant use and risk of adverse outcomes in older people: population based cohort study *BMJ* 2011;343:d4551
20. Shewale AR, Barnes CL, Fischbach LA, Ounpraseuth ST, Painter JT, Martin BC. Comparison of low-, moderate-, and high-molecular-weight hyaluronic acid injections in delaying time to knee surgery. *J Arthroplasty*. 2017;32(10):2952–2957

Table 1. Patient characteristics of new users of smoking cessation drugs (prior to trimming).

Characteristic	Varenicline (n = 68,468)	Bupropion (n = 9,931)	NRT (n = 155,339)
Age (years)			
Mean \pm SD	44.2 \pm 13.1	44.0 \pm 12.6	45.8 \pm 15.6
Median (min-max)	44.0 (18-106)	43.0 (18-88)	45.0 (18-108)
Women (n, %)	34,819 (50.9)	5,021 (50.6)	85,088 (54.8)
HES linkage available (n, %)	39,997 (58.4)	6,551 (66.0)	87,921 (56.6)
Year of cohort entry (n, %)			
2006	0 (0.0)	243 (2.4)	2,191 (1.4)
2007	10,056 (14.7)	5,015 (50.5)	44,792 (28.8)
2008	13,873 (20.3)	1,950 (19.6)	29,472 (19.0)
2009	12,919 (18.9)	1,317 (13.3)	27,391 (17.6)
2010	13,200 (19.3)	743 (7.5)	22,457 (14.5)
2011	11,868 (17.3)	481 (4.8)	18,642 (12.0)
2012	6,552 (9.6)	182 (1.8)	10,394 (6.7)
Body mass index (n, %)			
<30	41,832 (61.1)	6,075 (61.2)	96,115 (61.9)
\geq 30	12,886 (18.8)	1,766 (17.8)	30,369 (19.6)
Missing	13,750 (20.1)	2,090 (21.0)	28,855 (18.6)
Alcohol abuse (n, %)	7,259 (10.6)	896 (9.0)	21,667 (13.9)
Comorbidities (n, %)			
Coronary artery disease	4,447 (6.5)	576 (5.8)	13,209 (8.5)
Atrial fibrillation	475 (0.7)	57 (0.6)	2,091 (1.3)
Cerebrovascular disease	1,445 (2.1)	160 (1.6)	5,022 (3.2)
Congestive heart failure	336 (0.5)	41 (0.4)	1,559 (1.0)
Previous coronary revascularization	868 (1.3)	87 (0.9)	2,233 (1.4)
Previous myocardial infarction	1,244 (1.8)	146 (1.5)	3,657 (2.4)
Previous stroke	916 (1.3)	95 (1.0)	3,024 (1.9)
Hypertension	8,250 (12.0)	1,002 (10.1)	25,801 (16.6)
Peripheral vascular disease	2,216 (3.2)	299 (3.0)	6,952 (4.5)
Hypercholesterolemia	5,064 (7.4)	624 (6.3)	13,450 (8.7)
Chronic obstructive pulmonary disease	7,494 (10.9)	1,010 (10.2)	20,568 (13.2)
Diabetes	3,267 (4.8)	378 (3.8)	9,857 (6.3)
Depression	23,891 (34.9)	3,689 (37.1)	64,091 (41.3)
Medications in the year prior to cohort entry (n, %)			
Number of distinct drug class			
<4	29,506 (43.1)	4,531 (45.6)	50,213 (32.3)
4-6	18,666 (27.3)	2,851 (28.7)	42,316 (27.2)
7+	20,296 (29.6)	2,549 (25.7)	62,810 (40.4)
Aspirin	4,995 (7.3)	613 (6.2)	16,379 (10.5)
Angiotensin-converting enzyme inhibitors	5,387 (7.9)	614 (6.2)	16,016 (10.3)
Angiotensin receptor blockers	1,343 (2.0)	147 (1.5)	4,050 (2.6)
Beta-blockers	4,138 (6.0)	544 (5.5)	13,237 (8.5)

Calcium-channel blockers	4,016 (5.9)	451 (4.5)	12,858 (8.3)
Diuretics	3,955 (5.8)	531 (5.3)	14,247 (9.2)
Fibrates	226 (0.3)	29 (0.3)	644 (0.4)
Statins or statin combinations	9,099 (13.3)	1,089 (11.0)	25,880 (16.7)
Non-steroidal anti-inflammatory drugs	15,173 (22.2)	2,193 (22.1)	36,690 (23.6)

Abbreviations: HES: Hospital Episode Statistics; NRT: Nicotine replacement therapy; SD: Standard deviation.

Table 2. Association between smoking cessation therapy and rate of cardiovascular events using pairwise and multinomial HDPS modeling approaches.

Exposure	Number of Events/Patients	Total Number of Person- Years	Rate [*]	Hazard Ratio (95% CI)		
				Crude	Adjusted 1 – pre- specified only [†]	Adjusted 2 – pre- specified + empirically selected [§]
Pairwise Comparisons						
Varenicline	90/68,319	8,348.2	10.8	1.01	1.02 (0.54, 1.90)	1.08 (0.57, 2.04)
Bupropion	13/9,930	1,216.9	10.7	1.00	1.00 (Ref)	1.00 (Ref)
Varenicline	90/68,467	8,366.3	10.8	0.38	0.56 (0.45, 0.71)	0.64 (0.51, 0.81)
NRT	512/155,272	18,162.8	28.2	1.00	1.00 (Ref)	1.00 (Ref)
Bupropion	13/9,929	1,216.8	10.7	0.39	0.71 (0.41, 1.24)	0.76 (0.43, 1.33)
NRT	493/154,894	18,120.1	27.2	1.00	1.00 (Ref)	1.00 (Ref)
Multinomial HDPS						
Analysis 1 [€]						
Bupropion	13/9,924	1,216.0	10.7	0.38	0.65 (0.38, 1.14)	0.70 (0.40, 1.23)
Varenicline	90/68,414	8,359.8	10.8	0.38	0.55 (0.44, 0.70)	0.60 (0.48, 0.76)
NRT	507/155,059	18,139.4	28.0	1.00	1.00	1.00 (Ref)
Analysis 2 [£]						
Bupropion	13/9,930	1,216.9	10.7	0.39	0.66 (0.38, 1.15)	0.72 (0.41, 1.26)
Varenicline	90/68,301	8,345.0	10.8	0.39	0.57 (0.45, 0.72)	0.62 (0.49, 0.79)
NRT	488/154,232	18,040.6	27.1	1.00	1.00	1.00 (Ref)
Analysis 3 [¥]						

Bupropion	13/9,930	1,216.9	10.7	0.38	0.66 (0.38, 1.15)	0.70 (0.40, 1.23)
Varenicline	90/68,407	8,358.7	10.8	0.39	0.56 (0.44, 0.70)	0.62 (0.49, 0.78)
NRT	502/154,633	18,090.6	27.7	1.00	1.00	1.00 (Ref)

Abbreviations: CI, confidence interval; HDPS, high-dimensional propensity score; NRT, nicotine replacement therapy.

*Rates are expressed as events per 1,000 person-years.

† Propensity score model is based on predefined covariates only. In pairwise comparisons, adjusted for age, sex, propensity score deciles (9 dummy variables). In multinomial comparisons, adjusted for age, sex, propensity score quintiles (eight dummy variables; four for each of the two propensity scores).

§ Propensity score model is based on predefined covariates and the variables empirically selected by the HDPS algorithm. In pairwise comparisons, adjusted for age, sex, HDPS deciles (9 dummy variables). In multinomial comparisons, adjusted for age, sex, HDPS quintiles (eight dummy variables; four for each of the two propensity scores).

€Only the covariates that were among the top 500 in all three comparisons were used in the multinomial propensity score model, resulting in the adjustment for 263 empirically selected covariates.

£A total of 500 empirically selected covariates, including all concordant covariates and the next top ranked covariates from each comparison, were used in the multinomial propensity score model.

¥The top 500 covariates from the three comparisons were used in the multinomial propensity score model, resulting in the adjustment for 589 empirically selected covariates.

FIGURE LEGENDS

Figure 1. Construction of cohort of new users of smoking cessation drugs. Abbreviations:
NRT: Nicotine replacement therapy; Rx: Number of prescriptions.

Figure 1.

