Use of Propensity Score Technique to Account for Exposure-Related Covariates

An Example and Lesson

John D. Seeger, PharmD, DrPH,† Tobias Kurth, MD, ScD,†‡ and Alexander M. Walker, MD, DrPH*†*

Background: In observational research, propensity score techniques can be used to account for baseline differences between compared therapies. Although propensity scores are used increasingly often, their limitations in settings without complete data may not be recognized.

Objectives: We sought to evaluate the ability of propensity score matching to mitigate confounding by indication in an observational study of the effect of statin therapy on acute myocardial infarction (AMI). Matching was performed at random, and with propensity scores that incorporated a reduced or expanded set of variables.

Research Design/Subjects: This was a propensity score matched cohort study using members of a health insurer database.

Measures: Exposure to statin therapy was assessed at the beginning of follow-up with all cohort members being statin initiators or noninitiators, and the outcome of AMI was identified on the basis of claims codes.

Results: Matching on the basis of the propensity score provided results that are similar in magnitude to randomized clinical trials, suggesting that confounding was mitigated. However, matching on a propensity score created on a reduced set of variables yielded a result that suggested no effect of statin therapy, and demonstrated substantial imbalance on some variables that were not part of the propensity score.

Conclusions: Propensity score matching can balance with respect to variables not explicitly included in the score, but external data are required to evaluate this.

Key Words: cohort studies, propensity scores, observational research

(*Med Care* 2007;45: S143–S148)

From the *Ingenix i3 Drug Safety, Waltham, Massachusetts; †Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts; and ‡Divisions of Preventive Medicine and Aging, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts.

Supported by the Harvard School of Public Health Pharmacoepidemiology training grant.

Reprints: John D. Seeger, PharmD, DrPH, Ingenix i3 Drug Safety, 950 Winter Street, Suite 3800, Waltham, MA 02451. E-mail: john.seeger@i3drugsafety.com.

ISSN: 0025-7079/07/4500-0143

Drugs are prescribed to particular patients for a host of reasons. Chief among them is the indication for the drug—a patient must have the condition for which the drug is indicated. In addition, a collection of less explicit characteristics of both the patient and prescriber influence the choice of a particular drug for a particular patient at a particular time. The collection of these explicit and implicit characteristics will differentiate persons who receive a given therapy from those who receive a different drug for the same condition (and those who receive no drug), and if these characteristics are also prognostic of outcomes, then an unconfounded assessment of the drug's effect on those outcomes will depend on addressing the baseline differences between groups of patients that result from the prescribing process.

In a randomized drug trial, the allocation of subjects to treatment and control groups at random leads to groups that are similar with respect to both measured and unmeasured baseline characteristics. A formal way of saying this is that treatment allocation, being based on a random number, is uncorrelated with any possible patient characteristic. At least in expectation and in fact with large numbers, treatment allocation being uncorrelated with patient characteristics means that the distribution of characteristics is the same in all treatment groups. This baseline comparability supports the conclusion that any differences between groups in the occurrence of outcomes during follow-up must be due to the 1 characteristic that differs between the 2 groups by design: treatment.

In an observational setting, matching groups of patients (cohorts) on the characteristics that are part of the prescribing decision creates a balance with respect to measured characteristics that is even more complete than what results from randomization, so that treatment and control groups are identical at the start of follow-up. For example, if age, gender, pre-existing diagnoses, and concomitant drug use are the characteristics that lead to selection of statin therapy, then it may be possible to find both a statin-treated and untreated 47-year-old male with hypertension who is taking a calcium channel blocker. By forming treatment and untreated groups that are individually matched in this way, an observational study of drug effect is possible, given 2 considerations: (1) that the list of matching characteristics is complete (it includes all variables that actually went into the prescribing

decision so that no unmeasured predictor of treatment is present), and (2) that there is a control individual with each collection of attributes that can be matched to each individual who received the drug. In a real-world setting, this form of exhaustive matching is difficult because these considerations are in conflict. Many characteristics can plausibly enter into the prescribing decision, leading to such a large number of combinations of characteristics that it becomes impossible to find a control individual with exactly the same characteristics as each treated individual. For example, consider matching on age in 10 categories, gender in 2 categories, 5 prior diagnoses each in 2 categories, 5 prior drug therapies each in 2 categories, and preceding cost of care in 5 categories. The characteristics in this fairly simple example will involve 102,400 potential matching groups, so that finding an untreated individual for each treated individual will be almost impossible even within the context of extremely large data sets.

The propensity score addresses this "curse of dimensionality" by collapsing the multidimensional vector of pretreatment covariates into a single value that can be used to match individuals.[1–3] Matching on the propensity score has the effect of balancing all the variables that are components of the score, and thereby removing confounding from these variables when making comparisons between the matched groups.[4] The propensity score can also be used as a stratification variable, or for regression adjustment or weighting.

Statins inhibit endogenous cholesterol synthesis, and therapy with these drugs improves lipoprotein profiles in patients with dyslipidemia.[5–7] The primary lipoprotein effect of statin therapy is a reduction in serum concentrations of low-density lipoprotein (LDL), but they also decrease triglyceride (TG) and increase high-density lipoprotein (HDL) serum concentrations. Evidence that statin therapy prevents coronary heart disease events [including acute myocardial infarction (AMI)] comes directly from several large clinical end point studies.[8–13] These studies indicate that statin therapy reduces coronary heart disease events by 24–37% independently of baseline coronary heart disease and LDL. The National Cholesterol Education Program recommends drug therapy (often with a statin) for patients with hypercholesterolemia and other characteristics that represent risk factors for AMI.[14,15]

We conducted a propensity-score–matched cohort study of the effect of statin therapy on AMI where we built the propensity score twice: once using a restricted set of variables, and a second time using a broader set of variables. This work was an extension of work that has been published.[16,17] We sought to demonstrate what might happen when attempting to address confounding by indication using propensity scores in a setting where numerous randomized trials provide evidence that statins reduce the risk of AMI, strong a priori expectation of confounding exists, and where only a portion of the relevant variables for propensity score estimation were available or measured within the source data.

METHODS

Design and Data Source

Fallon Community Health Plan (FCHP) is a staff-model Health Maintenance Organization (based in Worcester, Massachusetts) with an enrollment in 1990 of approximately 110,000, and in 1998 of approximately 150,000. Members are covered for most outpatient and inpatient services including ambulatory care, approved specialists and outside referrals, laboratory work, and prescription drugs. Services provided to members are recorded in a computerized database with coded diagnoses and procedures allowing longitudinal evaluation of members' healthcare utilization.

We identified statin initiators and those eligible for statin initiation (based on having a physician visit within 6 months of an LDL test result) within half-year blocks of calendar time, and matched statin initiators to noninitiators serially across each of 9 calendar time blocks. Subjects were followed until disenrollment from the health plan or July 1999. The propensity score was estimated using unconditional logistic regression.

Pharmacy dispensations of statins were identified from pharmacy claims by national drug code supplemented by a text search for the drug name. Clinical covariates were defined as a claim for a medical service from a provider with a diagnosis coded according to International Classification of Diseases, 9th edition (ICD-9) codes. Two different ambulatory care claim dates with a code for a particular diagnosis were required before patients were assigned that diagnosis, whereas a single claim from an inpatient setting was required. Lipoprotein measures (LDL, HDL, and TG) were assumed to be valid for up to 6 months from the date of the laboratory test or until replaced by a new measurement unless statin therapy was initiated.

Identification of Patients and Outcomes

Members of FCHP with hypercholesterolemia (a recorded LDL >130 mg/dL) at any time between 1994 and 1998 were eligible for inclusion in this study. Eligible FCHP members who initiated therapy with a statin between the beginning of 1994 and the end of 1998 were matched (using propensity scores) to eligible FCHP members who did not initiate statin therapy. Matched cohorts were assembled within 9 half-year blocks of calendar time and an index date (date of first statin dispensation for statin initiators or a randomly chosen physician visit within the cohort accrual block for noninitiators) was assigned to each individual. Eligibility criteria for cohort entry (membership in FCHP for at least 1 year; LDL, TG, HDL levels all performed within the past 6 months; at least 1 physician visit in the cohort accrual block; no diagnosis of peripheral arterial disease; and not a current statin user) were all evaluated on the index date. (Fig. 1 is a schematic representation of the study.)

Propensity scores (predicted probability of statin initiation) were estimated for each FCHP member on their index date and used to match each statin initiator in a block to a contemporaneous noninitiator. Predicted probabilities of statin initiation (propensity scores) for each eligible member were estimated with unconditional logistic regression (1 model for each cohort accrual block) with outcome being initiation of statin within each half-year block of time, and predictors being derived from the history of claims preceding the index for each patient. This modeling approach allows the propensity score to flexibly account for temporal changes in drug selection. The propensity score model was developed twice: the first time

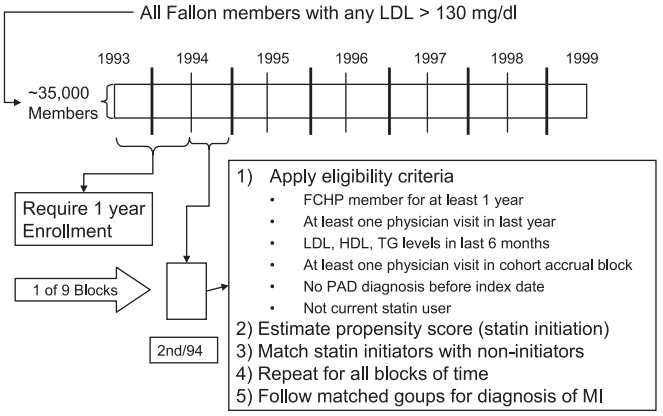

FIGURE 1. Study schematic.

including 38 variables and 4 quadratic terms chosen by selection based on review of literature relating to statin use and AMI risk factors, and the second time (after consulting with subject-matter experts in the field) including 52 variables and 6 quadratic terms. Within each cohort accrual block, each eligible statin initiator was matched to a noninitiator who had almost the same propensity score (within a 0.01 caliper). Initiators for whom no suitable noninitiator could be found were excluded from the matched cohort follow-up. Subjects in each cohort (statin initiator or noninitiator) were assigned their original exposure status until the end of follow-up regardless of actual statin use during follow-up, providing a conservative estimate of efficacy as intent-to-treat does in a clinical trial. The matched cohorts were followed from the date of matching through the end of follow-up to identify the occurrence of new-onset AMI.

The study outcomes (fatal or nonfatal AMI) were identified among the matched cohorts by the presence of an inpatient claim for AMI (ICD-9 code 410) received on 2 consecutive dates, or an inpatient claim for AMI on a single day followed by a complete cessation of claims (indicating fatal AMI). Survival analysis of time from date of matching until an event (first inpatient claim for AMI) was estimated using proportional hazards regression stratified by the cohort accrual block in which the matching had been performed. Subjects were censored on the occurrence of death, disenrollment, discontinuity of enrollment, or July 1999.

Analyses were conducted using SAS version 8.2 (SAS Institute, Inc., Cary, NC). This study was carried out under Institutional Review Board approvals from the Harvard School of Public Health and the FCHP.

RESULTS

There was an average of 504 (±60.0) statin initiators and 8090 (±1168) statin noninitiators among the FCHP members who were eligible for propensity score estimation and cohort inclusion in each half-year accrual block. Initial matching took all 4144 statin initiators and chose (at random) an equal number of noninitiators. The baseline characteristics of the subjects in these groups are quite different in clinically expected ways, with the statin initiators exhibiting more heart disease and heart disease risk factors, suggesting that the statin initiators are at increased risk of cardiovascular events (Table 1).

TABLE 1. Statin Initiators and Randomly* Selected Comparators

No.	Variable	Statin Initiators (N = 4144)	Noninitiators (N = 4144)	*P*
1	No. prescribed medications	5.0	2.9	<0.01
2	LDL (mg/dL)	180.3	155.1	<0.01
3	Triglycerides (mg/dL)	202.7	166.9	<0.01
4	Age (yr)	62.0	58.0	<0.01
5	No. physician visits in the past year	7.7	6.3	<0.01
6	Previously diagnosed ischemic heart disease	20.3%	5.6%	<0.01
7	HDL (mg/dL)	43.3	46.6	<0.01
8	Past acute myocardial infarction	11.6%	2.9%	<0.01
9	Diagnosis of angina	12.0%	3.1%	<0.01
10	Diagnosis of unstable angina	11.0%	2.2%	<0.01
11	History of smoking	25.8%	18.1%	<0.01
12	Diagnosis of hypertension	20.0%	13.0%	<0.01

*Statin noninitiators were randomly selected among eligible noninitiators within the same half-year calendar block as statin initiators.

TABLE 2. Statin Initiators and Matched Comparators (Reduced Model)

No.	Variable	Statin Initiators (N = 3579)	Noninitiators* (N = 3579)	*P*
1	No. prescribed medications	4.6	4.7	0.47
2	LDL (mg/dL)	177.5	177.4	0.90
3	Triglycerides (mg/dL)	197.4	198.1	0.82
4	Age (yr)	61.8	61.6	0.52
5	No. physician visits in the past year	7.4	7.5	0.51
6	Previously diagnosed ischemic heart disease	16.4%	14.8%	0.07
7	HDL (mg/dL)	43.7	43.7	0.89
8	Past acute myocardial infarction	8.7%	7.5%	0.08
9	Diagnosis of angina	9.3%	8.6%	0.34
10	Diagnosis of unstable angina	7.5%	6.9%	0.34
11	History of smoking	23.9%	24.9%	0.31
12	Diagnosis of hypertension	18.1%	18.0%	0.98

Reduced model contained 38 variables and 4 quadratic terms.
*Statin noninitiators were selected among eligible noninitiators within the same half-year calendar block as statin initiators provided they had sufficiently similar propensity scores (within 0.01).

The first set of propensity models (based on a reduced set of variables) was highly predictive of statin initiation (C-statistics ranged from 0.85 to 0.87, mean = 0.86), and it was possible to match 3579 of the 4144 statin initiators (86%) to noninitiators on the basis of this score (fewer subjects than among the randomly matched cohorts, because the statin initiators for whom no comparable noninitiator could be found were removed from follow-up). The characteristics of those who matched and their comparators are quite similar to one another on each of the variables tabulated (Table 2).

TABLE 3. Statin Initiators and Comparators (Expanded Model)

No.	Variable	Statin Initiators (N = 2901)	Noninitiators* (N = 2901)	*P*
1	No. prescribed medications	4.6	4.5	0.76
2	LDL (mg/dL)	177.8	177.6	0.78
3	Triglycerides (mg/dL)	200.3	200.5	0.96
4	Age (yr)	61.5	61.7	0.50
5	No. physician visits in the past year	7.2	7.3	0.87
6	Previously diagnosed ischemic heart disease	15.1%	15.5%	0.74
7	HDL (mg/dL)	43.5	43.5	0.91
8	Past acute myocardial infarction	7.9%	8.7%	0.32
9	Diagnosis of angina	8.5%	8.7%	0.82
10	Diagnosis of unstable angina	7.1%	7.3%	0.84
11	History of smoking	23.9%	24.3%	0.74
12	Diagnosis of hypertension	16.6%	18.0%	0.16

Expanded model included 52 variables and 6 quadratic terms.
*Statin noninitiators were selected among eligible noninitiators within the same half-year calendar block as statin initiators provided they had sufficiently similar propensity scores (within 0.01).

The second set of propensity models (based on an expanded set of variables) was even more predictive of statin initiation (C-statistics ranged from 0.90 to 0.94, mean = 0.92). It was possible to match 2901 of the 4144 statin initiators (70%) to noninitiators on the basis of this score. The characteristics of the matched subjects are again quite similar to one another (Table 3).

The characteristics of the reduced and expanded matched groups with respect to the variables that differ between the propensity scores are shown in Table 4. The matched groups based on the propensity score developed on the reduced set of variables exhibit close similarity with respect to some of the variables (such as counts of electrocardiograms or depression) and substantial difference with respect to other variables (such as lipid-related laboratory tests or cardiovascular diagnoses).

The numbers of AMI events among the matched groups of statin initiators and noninitiators along with hazard ratios are presented separately for the 3 sets of matched groups (randomly selected, reduced, and expanded sets of propensity score models; Table 5). Statin initiators have more AMIs than the randomly selected comparators, about the same number as the comparators selected on the basis of the propensity score developed using the reduced set of variables, and fewer than the comparators selected on the basis of the propensity score developed using the expanded set of variables.

CONCLUSIONS

We found a beneficial effect of statin therapy on the occurrence of AMI of similar magnitude to that seen in randomized clinical trials of lovastatin, pravastatin, and simvastatin.[4–9] However, this finding is dependent on the variables available for construction of the propensity score, and the finding could be quite different under different data situations with a null result under a reduced set of variables or

TABLE 4. Balance Between Statin Initiators and Noninitiators With Respect to the 14 Additional Variables Included in the Expanded Propensity Score Model

	Matched Groups Formed Using Propensity Score Based on					
	Reduced Set (38 Variables)			Expanded Set (52 Variables)		
Variable	Statin Initiators (N = 3579)	Noninitiators (N = 3579)	*P*	Statin Initiators (N = 2901)	Noninitiators (N = 2901)	*P*
Mean no. lipid lab tests in the past year	26.3	16.4	<0.01	24.9	24.6	0.50
Mean no. cardiovascular disease-related prescriptions	0.51	0.53	<0.09	0.51	0.51	0.94
Mean no. cardiovascular disease-related physician visits in the past year	0.96	0.54	<0.01	0.74	0.83	0.12
Mean no. cardiovascular disease-related diagnoses	0.25	0.17	<0.01	0.21	0.23	0.21
Mean no. cardiovascular disease-related hospitalizations in the past year	0.46	0.31	<0.01	0.40	0.39	0.79
Mean no. laboratory tests in the past year	10.3	11.3	<0.01	10.5	10.5	0.79
Mean no. electrocardiograms in the past year	0.48	0.48	0.77	0.48	0.51	0.35
Schizophrenia	2.2%	3.6%	0.03	2.1%	2.2%	0.99
Depression	2.4%	2.9%	0.14	1.9%	2.7%	0.05
Cancer (excluding nonmelanoma skin cancer)	5.3%	5.7%	0.47	5.2%	5.5%	0.60
Adjustment disorder	2.7%	3.6%	0.03	2.8%	3.1%	0.59
Skin cancer	2.5%	2.7%	0.71	2.2%	2.4%	0.73
Debility	1.7%	1.5%	0.78	1.6%	1.3%	0.44
Rheumatic disease	1.3%	1.8%	0.07	1.2%	1.7%	0.19

TABLE 5. ●●●

Matching	No. Cohort Members Diagnosed With an Acute Myocardial Infarction During Follow-up		Hazard Ratio	95% CI	*P*
	Statin Initiators	Statin Noninitiators			
Unmatched	325	124	2.11	1.46–3.04	<0.01
Reduced model propensity score	113	116	1.01	0.62–1.29	0.93
Expanded model propensity score	60	104	0.69	0.52–0.93	<0.01

a harmful effect if no variables are available. The result closer to the randomized trials was achieved through close matching on the propensity score, which had the effect of removing some statin initiators who could not be matched to a comparable noninitiator. This process improved the internal validity of the comparison, but it came at the cost of reduced generalizability. Other techniques such as weighting or modeling with the propensity score (or even widening the matching caliper) might allow for a larger fraction of the statin initiators to be included in the analysis so that results could be more broadly generalized, but their use may involve additional assumptions.

Propensity score matching can produce a high degree of balance on component variables, but may not achieve balance on variables that were not explicitly included in the propensity score. Health services utilization data (such as claims data) may allow for the formation of reasonable proxies for difficult to capture variables such as physician/patient belief about prognosis and risk.

Beyond its usual role as a variable reduction technique, propensity scoring can be thought of as an empirical method for creation of a multidimensional proxy variable that might represent an unmeasured variable. This consideration will have implications for variable selection within propensity scores, because these healthcare utilization variables may not have a priori expectations of association with the outcome variable.[18] Further, the degree to which these proxy variables represent an unmeasured variable is not known from the data and must be argued from external information such as clinical trial results, expert opinion, surveys/chart reviews, or sensitivity analyses. An efficient sampling scheme for validation of results based on calibration using external information has been developed.[19]

Observational methods such as we used in this study may offer an alternative for answering questions of efficacy in settings where randomized clinical trials would encounter practical or ethical barriers. However, observational studies of efficacy must address concerns of bias—especially confounding by indication where therapies are tailored to a patient's expected benefit, so statins are preferentially prescribed to patients at higher risk of AMI. This selective prescribing can overwhelm beneficial pharmacologic effects, leading to apparently higher risks of AMI among statin-treated persons. Appropriate use of propensity score methods can overcome this confounding. However, effect estimates from propensity score methods can be sensitive to modeling or variable selection and the apparent balance produced within observed variables can obscure a lack of balance on unobserved variables. In addition, appropriate selection of confounding control methods (with or without the propensity score) according to the question under study is required for causal inference.[20]

By matching patients according to propensity for treatment as estimated by a multivariate model, this process removes potential confounding by any variable that is part of the propensity model. Thus, a model that appropriately adjusts for all potential confounding variables will create closely matched cohorts from which to obtain unbiased effect estimates. However, variables that are not part of the propensity score could remain unbalanced between cohorts, and the residual confounding effect of these unbalanced variables might lead to a biased assessment of treatment effects.

We found that measures of healthcare utilization (intensities of lipid-related laboratory testing, medication use, and physician visits) were important predictors of statin initiation and their inclusion in the propensity score model resulted in greater control of confounding than a model that did not include them. Healthcare seeking behavior and willingness to begin a new chronic therapy as well as the physician's urgency in recommending a therapy will be reflected in these measures to the extent that they alter a patient's pattern of healthcare utilization.

REFERENCES

1. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70:41–55.
2. Joffe MM, Rosenbaum PR. Invited commentary: propensity scores. *Am J Epidemiol*. 1999;150:327–333.
3. Rubin DB. Estimating causal effects from large data sets using propensity scores. *Ann Intern Med*. 1997;127:757–763.
4. Rubin DB, Thomas N. Matching using estimated propensity scores: relating theory to practice. *Biometrics*. 1996;52:249–264.
5. Knopp RH. Drug treatment of lipid disorders. *N Engl J Med*. 1999;341:498–511.
6. Maron DJ, Fazio S, Linton MRF. Current perspectives on statins. *Circulation*. 2000;101:207–213.
7. Farnier M, Davignon J. Current and future treatment of hyperlipidemia: the role of statins. *Am J Cardiol*. 1998;82(Suppl J):3J–10J.
8. Scandinavian Simvastatin Survival Study Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet*. 1994;344:1383–1389.
9. Shepherd J, Cobbe SM, Ford I, et al; West of Scotland Coronary Prevention Study Group. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. *N Engl J Med*. 1995;333:1301–1307.

10. Sacks FM, Pfeffer MA, Moye LA, et al; Cholesterol and Recurrent Events Trial Investigators. The effect of pravastatin on coronary events after myocardial infarction in patients with average cholesterol levels. *N Engl J Med*. 1996;335:1001–1009.
11. Downs JR, Clearfield M, Weis S, et al; Air Force/Texas Coronary Atherosclerosis Prevention Study. Primary prevention of acute coronary events with lovastatin in men and women with average cholesterol levels: results of AFCAPS/TexCAPS. *JAMA*. 1998;279:1615–1622.
12. The Long-Term Intervention with Pravastatin in Ischaemic Disease (LIPID) Study Group. Prevention of cardiovascular events and death with pravastatin in patients with coronary heart disease and a broad range of initial cholesterol levels. *N Engl J Med*. 1998;339:1349–1357.
13. Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high-risk individuals: a randomised placebo-controlled trial. *Lancet*. 2002;360:7–22.
14. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel II). Summary of the second report of the National Cholesterol Education Program (NCEP). *JAMA*. 1993;269:3015–3023.
15. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Executive summary of the third report of the National Cholesterol Education Program (NCEP). *JAMA*. 2001;285:2486–2497.
16. Seeger JD, Walker AM, Williams PL, et al. A propensity score-matched cohort study of the effect of statin, mainly fluvastatin on the occurrence of acute myocardial infarction. *Am J Cardiol*. 2003;92:1447–1451.
17. Seeger JD, Williams PL, Walker AM. An application of propensity score matching using claims data. *Pharmacoepidemiol Drug Saf*. 2005;14:465–476.
18. Brookhart MA, Schneeweiss S, Rothman KJ, et al. Variable selection for propensity score models. *Am J Epidemiol*. 2006;163:1149–1156.
19. Sturmer T, Schneeweiss S, Avorn J, et al. Adjusting effect estimates for unmeasured confounding with validation data using propensity score calibration. *Am J Epidemiol*. 2005;162:279–289.
20. Kurth T, Walker AM, Glynn RJ, et al. Results of multivariable logistic regression, propensity matching, propensity adjustment, and propensity-based weighting under conditions of non-uniform effect. *Am J Epidemiol*. 2006;163:262–270.