Practice of Epidemiology

Reducing the Variance of the Prescribing Preference-based Instrumental Variable Estimates of the Treatment Effect

Michal Abrahamowicz*, Marie-Eve Beauchamp, Raluca Ionescu-Ittu, Joseph A. C. Delaney, and Louise Pilote

* Correspondence to Dr. Michal Abrahamowicz, Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Royal Victoria Hospital, 687 Pine Avenue West, V2.20A, Montreal, Quebec, Canada H3A 1A1, (e-mail: michal.abrahamowicz@mcgill.ca).

Instrumental variable (IV) methods based on the physician's prescribing preference may remove bias due to unobserved confounding in pharmacoepidemiologic studies. However, IV estimates, originally defined as the treatment prescribed for a single previous patient of a given physician, show important variance inflation. The authors proposed and validated in simulations a new method to reduce the variance of IV estimates even when physicians' preferences change over time. First, a potential "change-time," after which the physician's preference has changed, was estimated for each physician. Next, all patients of a given physician were divided into 2 homogeneous subsets: those treated before the change-time versus those treated after the change-time. The new IV was defined as the proportion of all previous patients in a corresponding homogeneous subset who were prescribed a specific drug. In simulations, all alternative IV estimators avoided strong bias of the conventional estimates. The change-time method reduced the standard deviation of the estimates by approximately 30% relative to the original previous patient-based IV. In an empirical example, the proposed IV correlated better with the actual treatment and yielded smaller standard errors than alternative IV estimators. Therefore, the new method improved the overall accuracy of IV estimates in studies with unobserved confounding and time-varying prescribing preferences.

confounding factors (epidemiology); decision support techniques; pharmacoepidemiology

The study of pharmacoepidemiology increasingly relies on large administrative health databases (1, 2) in which many clinical risk factors are not recorded (3, 4), which can lead to concerns about unobserved confounding (5). One generic approach that, under certain assumptions, may remove unobserved confounding bias uses instrumental variables (IVs), which are associated with the outcome exclusively through the treatment (6–8). However, the variance of estimates based on weak instruments may be inflated to the extent that the estimates may be less accurate than the biased conventional estimates (8, 9). Yet, many instruments used in observational studies of medications

are relatively weak (10, 11). Therefore, it is important to develop and validate methods to increase the strength of instruments and stabilize the IV estimates.

Although finding an adequate instrument in pharmacoepidemiology is challenging (12), in 2006, Brookhart et al. (13) proposed to define the IV as the physician's subjective preference for 1 of 2 alternative treatments. Indeed, important differences between individual physicians' prescribing patterns persist independently of patients' characteristics (14, 15). Because subjective preferences are not observable, the IV was operationalized as the treatment prescribed to the previous patient of the same physician (13). The IV was

based on the treatment prescribed to a single previous patient because of concerns about possible changes over time in physicians' preferences (13), as documented by other empirical studies (10, 15). Such changes may be triggered by publication of the results of influential clinical trials (16, 17), visits from pharmaceutical company representatives (18), "Dear Doctor" letters (19), or a recent negative clinical experience with a previously preferred drug. Some of these factors could induce an abrupt change in a short time interval (20–22).

On the other hand, defining the IV on the basis of only a single (previous) patient may inflate the variance of the IV estimates by 1) dichotomizing the instrument and 2) making it very dependent on the characteristics of the previous patient (9). To address these concerns, we proposed and validated in simulations a new method with which we attempted to stabilize the physician preference-based IV estimates when these preferences do change over time.

MATERIALS AND METHODS

The physician preference-based IV estimator by Brookhart et al.

Brookhart et al. (13) operationally defined the IV as the treatment prescribed to the single most recent patient of a given physician. First, the n_i patients of a given ith physician are ordered by the increasing date of treatment initiation, and then the IV for patient $j = 2,\ldots, n_i$ is defined as the treatment prescribed to patient $j − 1$. The IV approach was implemented by using 2-stage least-squares regression, which is recommended for weak instruments (6, 23, 24). The 2-stage least-squares analysis involves estimating 2 consecutive multiple linear regression models (23). The first model predicts the probability of patient j ($j = 2,\ldots, n_i$) of a given physician i being prescribed drug "B" ($T_{ij} = 1$ vs. $T_{ij} = 0$ for drug "A"), conditional on the instrument (IV_{ij}) and the vector X_{ij} of m covariates ($X_{1ij} − X_{mij}$) of patient j of physician i:

$$P(T_{ij} = 1|\text{IV}_{ij}, \mathbf{X}_{ij}) = \beta_0 + \beta_1 \text{IV}_{ij} + \sum_{p=1}^{m} \beta_{(p+1)} X_{pij}, \quad (1)$$

where $\text{IV}_{ij} = T_{i(j − 1)}$ (i.e., the treatment assigned to the previous patient $j − 1$ of the same physician) and $\hat{\beta}_1$ quantifies the association between the IV and the actual treatment in terms of adjusted risk difference.

The second model predicts the outcome of the jth patient (Y_{ij}) conditional on probability of treatment with drug B, estimated in the first step ($P(T_{ij} = 1|\text{IV}_{ij}, X_{ij})$), and observed covariates ($X_{1ij} − X_{mij}$):

$$P(Y_{ij} = 1|P(T_{ij} = 1|\text{IV}_{ij}, \mathbf{X}_{ij}), \mathbf{X}_{ij})$$
$$= \alpha_0 + \alpha_1 P(T_{ij} = 1|\text{IV}_{ij}, \mathbf{X}_{ij}) + \sum_{p=1}^{m} \alpha_{(p+1)} X_{pij}, \quad (2)$$

where $\hat{\alpha}_1$ is the IV estimate of the effect (adjusted risk difference) of the treatment on the outcome (23).

New "change-time" IV estimator

We proposed a new method that modifies the IV approach originally proposed by Brookhart et al. (13) to measure a physician's preference based on prescriptions dispensed to a larger number of patients while accounting for potential changes over time in physicians' preferences. To this end, we first adapted the change-point method (25) to test whether the physician's prescribing preference did change during the study period and estimated the time of such change.

The algorithm involves the following 5 steps, which are repeated for each of $i = 1,\ldots, N$ physicians in the database:

1. Rank the n_i patients who received prescriptions from the ith physician by increasing calendar time of their prescriptions.
2. Fit a multivariable logistic regression no-change model, which predicts the probability of the ith physician's patients being prescribed drug B ($T_{ij} = 1$) based on available patients' characteristics $X_{1ij} − X_{mij}$, while assuming the preference remained constant across the study period:

$$\text{Logit}[P(T_{ij} = 1)] = \beta_0 + \sum_{p=1}^{m} \beta_p X_{pij}. \quad (3)$$

Denote by $D(0)_i$ the deviance of the no-change model (equation 3) for physician i.
3. Test whether there is any systematic change in the prescribing preference of physician i and, if so, estimate the time of this change. This step involves 3 substeps:
   3.1. For each patient ($j = 3,\ldots, n_i − 3$) of physician i, estimate the difference (d_{ij}) in the proportions of prescriptions for drug B among "later" ($k > j$) versus "earlier" ($k \leq j$) patients (changes affecting <4 patients cannot be reliably identified):

$$d_{ij} = \frac{\sum_{1 \leq k \leq j} T_{ik}}{j} − \frac{\sum_{j+1 \leq k \leq n_i} T_{ik}}{n_i − j}. \quad (4)$$

   If $|d_{ij}| < 0.2$ for all patients j, we assume there is no change in physician i's preference and thus omit steps 3.2, 3.3, and 4. However, if $|d_{ij}| \geq 0.2$ for at least 1 j ($3 < j < n_i − 3$), we assume physician i's preference did change and proceed to estimate the time of such change in substeps 3.2 and 3.3.
   3.2. Fit for each patient j for whom $|d_{ij}| \geq 0.2$ a change-time model that expands the no-change model (equation 3) by adding a binary indicator of the patients being prescribed after the jth patient (i.e., $I\{k > j\} = 1$ if $k > j$ or $I\{k > j\} = 0$ for $k \leq j$):

$$\text{Logit}[P(T_{ij} = 1)]$$
$$= \beta_0 + \sum_{p=1}^{m} \beta_p X_{pij} + \gamma I\{k > j\}. \quad (5)$$

The coefficient γ in equation 5 represents the average adjusted difference in the ith physician's propensity to prescribe drug B between later ($k > j$) and

earlier ($k \leq j$) patients. Denote by $D_i(j)$ the deviance of the change-time model (equation 5), with the presumed change after the jth patient.

3.3. Identify the optimal change-time model that corresponds to the lowest deviance among all patient-specific models estimated in step 3.2 and denote its deviance as $D_i(j^*)$ and the corresponding patient's index by j^*.

4. Compare the fit of the no-change model (equation 3) from step 2 with the optimal change-time model (equation 5) from step 3.3, based on the Akaike information criterion, calculated as deviance $+ 2 \times P$, where P is the number of parameters in a model (26). Given that the optimal model in equation 5 involves estimating 2 additional parameters with respect to the model in equation 3 (i.e., the optimal change-time j^* in step 3.3 and the corresponding change in preference γ in equation 5), the fit of the no-change model (equation 3) is considered better if and only if

$$D(0)_i < D_i(j^*) + 4, \qquad (6)$$

that is, if its deviance is less than 4 points higher than that of the optimal change-time model.

5. Define the change-time model IV* estimator:

5.1. If in step 3.1 $|d_{ij}| < 0.2$ for all $j = 3,\ldots, n_i - 3$ or in step 4 $D(0)_i < D_i(j^*) + 4$, we consider the preference of physician i to be constant over time. Accordingly, for each patient $j = 2,\ldots, n_i$, we define the new IV$_{ik}$* as the proportion of all previous patients $(1,\ldots, j-1)$ of physician i who were prescribed drug B.

5.2. If $|d_{ij}| \geq 0.2$ in step 3.1 and $D(0)_i \geq D_i(j^*) + 4$ in step 4, we conclude that the physician's preference changed between patients with indexes j^* and $j^* + 1$. Then, for patients seen after the estimated change $(j > j^*)$, the new IV$_{ij}$* is calculated as the proportion of patients prescribed drug B among previous patients seen after the change $(j^* + 1,\ldots, j-1)$. For patients with $j < j^*$, we use all previous patients $(k = 1,\ldots, j-1)$.

An SAS program (SAS Institute, Inc., Cary, North Carolina) containing an example that implements the method is available upon request.

SIMULATION DESIGN

We simulated a hypothetical database study comparing the risks of a binary adverse event between 2 drugs (A vs. B) in the presence of strong unmeasured confounding. The prescribing physician's identification number and the date of each prescription were recorded. Initially, most physicians preferred drug A over drug B, but some physicians switched their preferences at a random time during the study.

The probability of a given patient's being prescribed drug B depended on both the physician's current (at the time of the prescription) subjective prescribing preference and the patient's characteristics, which were represented by 2 continuous variables ($X1$ and $X3$) and 1 binary ($X2$) variable. All 3 patient's characteristics were also associated with the

outcome. We assumed that $X1$ and $X2$ were measured and adjusted for in all analyses, whereas $X3$ was not recorded and thus acted as the unobserved confounder (27). The binary outcomes were generated from risk difference (linear) models. We assumed that all other assumptions underlying the IV method (6, 8, 13, 23) were met.

Details of data generation are presented in Web Appendix 1 (available at http://aje.oxfordjournals.org/). In summary, we generated 200 physicians, among whom 40% initially preferred drug B. In different simulated scenarios, the number of patients per physician varied from 10 to 50 or from 50 to 200. The frequency of preference switches among physicians who initially preferred drug A versus drug B varied across scenarios: 1) 33% versus 10%, 2) 50% versus 20%, and 3) 75% versus 20%. We generated from a uniform distribution the index of the patient after whom the physician's preference changed.

For each patient, the probability of being prescribed drug B ($T_{ji} = 1$) was derived from a multivariable linear model and depended on the 3 covariates and the current physician's preference. Physicians' preferences were quantified by an increase in the probability of a patient's receiving drug B if his or her physician currently preferred drug B. We used as strength of preference for drug B (b_{PPij}) either $\beta_{PPij} = 0.70$ for all physicians or we generated individual β_{PPij} values from uniform distribution $U(0.5, 0.9)$. Individual binary outcomes were generated based on the probability of $Y_{ji} = 1$ derived from a linear model, conditional on the actual treatment and covariates. The parameters used in different simulated scenarios are listed in the first 2 columns of Tables 1–3.

The last scenario (scenario 13) was similar to scenario 12 except it assumed smooth changes over time in prescribing preferences of individual physicians. Specifically, we assumed a 3-linear model of changes in preferences of physicians assigned to the changing preference subgroup. The simulated patterns of changes varied across individual physicians in this subgroup depending on 3 randomly assigned parameters: 1) time (i.e., patient number) s_i, after which the preference started to change, 2) time interval l_i, over which the preference changed, and 3) the final strength P_i of the preference for drug B at the end of interval l_i. Accordingly, physician i, who switched from drug A to drug B after patient s_i, was assumed to: 1) have no preference ($P_{ij} = 0$) for drug B for his/her patients with indexes $j \leq s_i$, 2) have linearly increasing strength of preference for patients with indexes $s_i < j < (s_i + l_i)$, with $\beta_{PPij} = P_i[(j - s_i)/l_i]$, and 3) have the final preference of P_i for all patients with indexes $j \geq (s_i + l_i)$. For physicians who switched from B to A, the preferences changed in the opposite direction but followed the same patterns. Parameters s_i and l_i were generated so that the change of physician i's preference might have started before the first patient seen in the study period ($s_i < 1$) and/or ended after the last patient ($[s_i + l_i] > n_i$). See Web Appendix 1 for details.

Analysis of simulated data

For each simulated scenario, we generated 1,000 independent random samples. Each sample was analyzed using 5 models that were adjusted for the treatment effect of $X1$ and $X2$ but not for the unobserved confounder $X3$. In

Table 1. Relative Bias (%) of Treatment Effect Estimates Yielded by the Change-Time Model and the 4 Alternative Models

	Simulation Scenario				Relative Bias, %				
Scenario No.	No. of Patients per Physician	Physicians Whose Preference Switched From A to B, %	Physicians Whose Preference Switched From B to A, %	Preference for Drug B[b]	Change-Time Instrumental Variable Model	Alternative Model[a]			
						Model 1	Model 2	Model 3	Model 4
1	10–50	33	10	0.7	5.6	51.6	0.7	2.5	1.9
2	10–50	33	10	$U(0.5, 0.9)$	6.0	48.3	2.8	0.9	1.2
3	10–50	50	20	0.7	4.6	50.2	−1.8	−0.5	−0.1
4	10–50	50	20	$U(0.5, 0.9)$	6.9	49.4	2.3	2.6	2.1
5	10–50	75	20	0.7	12.9	53.8	5.7	7.0	7.1
6	10–50	75	20	$U(0.5, 0.9)$	5.9	46.7	4.7	0.9	0.4
7	50–200	33	10	0.7	−0.7	50.2	−2.6	−2.5	−2.5
8	50–200	33	10	$U(0.5, 0.9)$	1.5	46.4	0.4	0.6	0.2
9	50–200	50	20	0.7	4.6	51.7	4.1	3.6	3.1
10	50–200	50	20	$U(0.5, 0.9)$	3.2	47.5	1.3	1.3	1.9
11	50–200	75	20	0.7	0.7	49.7	−3.8	−0.7	−1.1
12	50–200	75	20	$U(0.5, 0.9)$	2.5	48.0	0.6	2.3	1.0
13	50–200	75	20	Smooth change	3.4	47.9	−2.1	1.1	

[a] Model 1 was the conventional model; in model 2, the instrumental variable was based on the last patient; in model 3, the instrumental variable was based on all previous patients; and in model 4, the instrumental variable was based on the true change-time.

[b] Increase in the probability of prescribing drug B for patients of physicians preferring this drug: 0.7 = fixed risk difference (70%) for all physicians; $U(0.5, 0.9)$ = risk difference generated from a uniform distribution $U(0.5, 0.9)$; and smooth change = smooth change in preference.

addition to the proposed change-time IV model, the following models were estimated:

1. model 1, the conventional model that was fitted through ordinary least squares;
2. model 2, which used the last patient-based IV proposed by Brookhart et al. (13);
3. model 3, which defined the IV as the proportion of all previous patients of physician i who were prescribed drug B (i.e., ignoring any potential changes in the physicians' preferences); and
4. model 4, the unrealistic IV model, which was similar to the change-time model but was based on the "true" (simulated) change-time for each physician.

On the basis of results aggregated across simulated samples, alternative treatment effect estimates (risk differences) were compared with respect to relative bias, empirical standard deviation, overall accuracy (quantified by the root mean squared error (RMSE)) and the coverage rate of the 95% confidence intervals. The strengths of the alternative IVs were assessed through their partial correlation with the actual treatments of individual patients, based on the first stage 2-stage least squares model. See Web Appendix 2 for details.

Simulations were performed using SAS, version 9.2.

RESULTS

Relative bias and coverage rates

Table 1 shows the relative bias of the treatment-effect estimates obtained by using alternative models. Because

of failure to adjust for the unobserved confounder, conventional model 1 underestimated the treatment effect by about 50% (Table 1) and yielded coverage rates of the 95% confidence intervals as low as 51%–85% or 42%–50% for scenarios with 10–50 or 50–200 patients per physician, respectively (data not shown). In contrast, all IV-based models (change-time model and models 2–4) produced relative biases uniformly below ±15% (Table 1) and accurate coverage rates between 94% and 96% (data not shown). The bias of change-time model estimates was somewhat larger than that for simpler IV models, even if the 95% confidence intervals for bias of different IV estimates overlapped (data not shown).

Empirical standard deviations

Table 2 shows the standard deviations of change-time model estimates (sixth column) and their ratios to the standard deviations of the 4 other models (ratios <1 in columns 7–10 indicate lower variance for change-time model estimates). Conventional model 1 produced systematically lower standard deviations. There were systematic, marked differences in standard deviations of alternative IV estimates. The change-time model reduced by about 30% the standard deviation of model 2 (IV based on the last patient) (13) and by 6%–21% the standard deviation of model 3 (IV based on all previous patients). The latter reduction increased with increasing frequency of changes in physicians' preferences, which were ignored in model 3. Interestingly, the variance reduction, relative to last patient-based IV model 2, achieved by the change-time model in scenario 13, with smooth gradual changes in physician preferences,

Table 2. Empirical Standard Deviations of Treatment Effect Estimates Yielded by the Change-Time Model Relative to Standard Deviations Yielded by the 4 Alternative Models

	Simulation Scenario				SD of Estimates From the Change-Time IV Model	Ratio of SDs[a,b]			
Scenario No.	No. of Patients per Physician	Physicians Whose Preference Switched From A to B, %	Physicians Whose Preference Switched From B to A, %	Preference for Drug B[c]		Model 1	Model 2	Model 3	Model 4
1	10–50	33	10	0.7	0.0155	1.53	0.73	0.94	1.00
2	10–50	33	10	$U(0.5, 0.9)$	0.0148	1.57	0.77	0.90	0.98
3	10–50	50	20	0.7	0.0149	1.55	0.72	0.83	0.98
4	10–50	50	20	$U(0.5, 0.9)$	0.0150	1.51	0.75	0.86	0.98
5	10–50	75	20	0.7	0.0154	1.63	0.71	0.79	0.99
6	10–50	75	20	$U(0.5, 0.9)$	0.0150	1.54	0.72	0.85	1.00
7	50–200	33	10	0.7	0.0070	1.47	0.68	0.91	1.00
8	50–200	33	10	$U(0.5, 0.9)$	0.0069	1.42	0.73	0.93	1.00
9	50–200	50	20	0.7	0.0069	1.45	0.70	0.89	1.00
10	50–200	50	20	$U(0.5, 0.9)$	0.0072	1.50	0.72	0.87	1.00
11	50–200	75	20	0.7	0.0070	1.50	0.69	0.83	0.99
12	50–200	75	20	$U(0.5, 0.9)$	0.0070	1.46	0.71	0.82	0.99
13	50–200	75	20	Smooth change	0.0083	1.75	0.65	0.94	

Abbreviations: IV, instrumental variable; SD, standard deviation.

[a] Ratio of the change-time model-based SD (shown in column 6) to the SD of the respective alternative model.

[b] Model 1 was the conventional model; in model 2, the instrumental variable was based on the last patient; in model 3, the instrumental variable was based on all previous patients; and in model 4, the instrumental variable was based on the true change-time.

[c] Increase in the probability of prescribing drug B for patients of physicians preferring this drug: 0.7 = fixed risk difference (70%) for all physicians; $U(0.5, 0.9)$ = risk difference generated from a uniform distribution $U(0.5, 0.9)$; and smooth change = smooth change in preference.

was similar to scenario 12, with abrupt changes (Table 2). Finally, change-time model estimates had standard deviations similar to unrealistic model 4, which relied on the true time of change.

Overall accuracy

Figure 1 shows the overall accuracy of the treatment effect estimates across the 13 scenarios described in Tables 1–3, with a higher RMSE indicating lower accuracy. Higher RMSEs for scenarios 1–6, in which there were 10–50 patients per physician, reflect the lower sample size. In those scenarios, conventional model 1 produced RMSEs that were slightly but systematically lower than those from the change-time model. However, in scenarios 7–13, which had a larger sample size, change-time model estimates had higher overall accuracy (Figure 1). Comparisons of RMSEs of practically unbiased IV estimates mostly reflect the differences in their standard deviations, as shown in Table 2. Accordingly, the proposed change-time model estimates had systematically higher overall accuracy than did the alternative IV models 2 and 3, including scenario 13, which had gradual changes in preferences (Figure 1). Still, when comparing the 3 realistic IV estimates in individual samples, the change-time model estimate was closer to the true risk difference in 37%–44% of samples, versus 24%–31% for last patient-based model 2 and 29%–35% for all patients-based model 3 (Web Table 1).

Strength of instruments

Table 3 shows the mean squared partial correlation between alternative IVs and actual treatment. The change-time model yielded, on average, a stronger instrument than did models 2 and 3 (Table 3), which explains the lower variance of change-time model estimates (Table 2). Web Appendix 3 provides results on the accuracy of detection of preference changes and their estimated times in the change-time model.

APPLICATION

To compare the performance of alternative IV estimators in a real-life database study, we reassessed the comparative effectiveness of rhythm versus rate control therapy in preventing death within 3 years in patients diagnosed with atrial fibrillation in 1999–2004. In 2002, the Atrial Fibrillation Follow-Up Investigation of Rhythm Management (AFFIRM) (28) reported no difference in mortality rates between the 2 treatments, resulting in a significant decrease in the prescribing of rhythm control drugs. Thus, we expected frequent changes of treatment preferences during the study period.

We used population-based data from the province of Quebec, Canada. The study population included elderly patients at least 66 years of age who had been hospitalized with a diagnosis of atrial fibrillation from 1999 to 2004 and
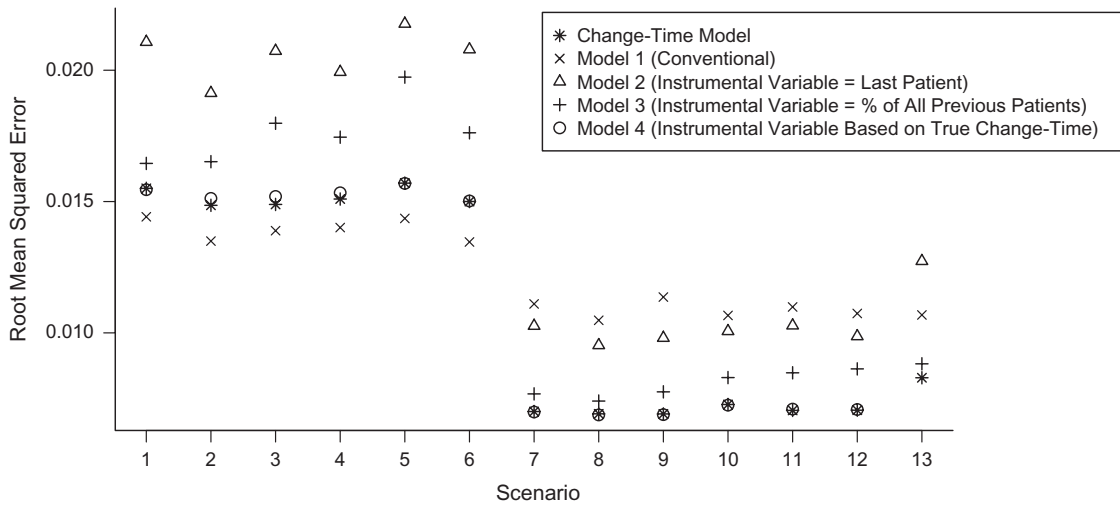
Figure 1.   Root mean squared errors of treatment effect estimates yielded by the 5 models. See Tables 1–3 for descriptions of the 13 scenarios.

who had received the first atrial fibrillation prescription within 7 days of discharge. Individual prescription records were linked with claims records for medical services, hospital discharge records, and mortality records. All subjects had 3 years of follow-up after their initial atrial fibrillation prescription after discharge, and the binary outcome was

death within those 3 years. The binary treatment variable was defined as either rhythm or rate control therapy, based on the first prescription after discharge, using the same drugs as in AFFIRM (28). Web Appendix 4 provides details on the study population and covariates for which we adjusted in all analyses. The severity of atrial fibrillation

Table 3.   Mean Squared Partial Correlation Between the Instrument and the Actual Treatment

	Simulation Scenario				Mean Squared Partial Correlation[a]			
Scenario No.	No. of Patients per Physician	Physicians Whose Preference Switched From A to B, %	Physicians Whose Preference Switched From B to A, %	Preference for Drug B[c]	Change-Time Instrumental Variable Model	Alternative Model[b]		
						Model 2	Model 3	Model 4
1	10–50	33	10	0.7	0.43	0.23	0.35	0.42
2	10–50	33	10	$U(0.5, 0.9)$	0.45	0.25	0.37	0.44
3	10–50	50	20	0.7	0.43	0.23	0.31	0.42
4	10–50	50	20	$U(0.5, 0.9)$	0.45	0.24	0.33	0.44
5	10–50	75	20	0.7	0.42	0.21	0.26	0.40
6	10–50	75	20	$U(0.5, 0.9)$	0.44	0.23	0.28	0.43
7	50–200	33	10	0.7	0.47	0.23	0.39	0.46
8	50–200	33	10	$U(0.5, 0.9)$	0.49	0.25	0.41	0.48
9	50–200	50	20	0.7	0.47	0.24	0.35	0.46
10	50–200	50	20	$U(0.5, 0.9)$	0.49	0.25	0.37	0.48
11	50–200	75	20	0.7	0.46	0.23	0.31	0.45
12	50–200	75	20	$U(0.5, 0.9)$	0.48	0.25	0.32	0.47
13	50–200	75	20	Smooth change	0.38	0.16	0.33	

[a] Mean value of the squared partial correlation, estimated from the first stage of the 2-stage least squares model (see equation 1), between the IV and the actual treatments.

[b] Model 1 was the conventional model; in model 2, the instrumental variable was based on the last patient; in model 3, the instrumental variable was based on all previous patients; and in model 4, the instrumental variable was based on the true change-time.

[c] Increase in the probability of prescribing drug B for patients of physicians preferring this drug: 0.7 = fixed risk difference (70%) for all physicians; $U(0.5, 0.9)$ = risk difference generated from a uniform distribution $U(0.5, 0.9)$; and smooth change = smooth change in preference.

Am J Epidemiol.   2011;174(4):494–502

Table 4. Risk Differences Between 3-Year Mortality Rates With Rhythm and Rate Control Therapy in Atrial Fibrillation Patients

Model[a]	Adjusted Risk Difference	Standard Error	95% Confidence Interval	P Value	Partial Correlation Between the Instrumental Variable and Treatment
Analyses adjusted for all important measured confounders					
Change-time IV model	0.0278[b]	0.0396	−0.0498; 0.1053	0.4832	0.1901
Model 1	−0.0026[b]	0.0075	−0.0173; 0.0121	0.7559	
Model 2	0.0967[b]	0.1159	−0.1315; 0.3228	0.4092	0.0648
Model 3	0.0652[b]	0.0509	−0.0345; 0.1648	0.2002	0.1477
Analyses not adjusted for 2 important measured confounders					
Change-time IV model	−0.0088[c]	0.0400	−0.0872; 0.0696	0.8254	0.1918
Model 1	−0.0522[c]	0.0076	−0.0671; −0.0372	<0.0001	
Model 2	0.0612[c]	0.1191	−0.1723; 0.2947	0.6074	0.0643
Model 3	−0.0069[c]	0.0509	−0.1066; 0.0927	0.8914	0.1505

Abbreviation: IV, instrumental variable.

[a] Model 1 was the conventional model; in model 2, the instrumental variable was based on the last patient; and in model 3, the instrumental variable was based on all previous patients.

[b] With adjustment for age, sex, history of diabetes, history of congestive heart failure, type of atrial fibrillation diagnosis at admission, length of stay, and physician specialty.

[c] Adjusted for sex, history of diabetes, history of congestive heart failure, length of stay, and physician specialty.

symptoms, which is an important predictor of mortality (29, 30), was not recorded, thereby raising concerns about potential unmeasured confounding.

Data were analyzed using the conventional multivariable linear regression model and the 3 IV models utilized in our simulations. Similar to what was done in other studies (31, 32), the IVs were defined at the level of the hospitals rather than the physicians. Indeed, the postdischarge atrial fibrillation treatment was likely initiated in the hospital. Furthermore, hospitals are an important exogenous source of variation in drug prescriptions (33, 34), as physicians within the same hospital often follow treatment guidelines set by local opinion leaders (35). Finally, in preliminary analyses, hospital-level instruments correlated much better with the actual treatment than the physician-based instruments (data not shown).

Among 19,688 atrial fibrillation patients discharged from 89 hospitals, 7,698 deaths occurred within 3 years of discharge. A change in prescribing preferences was detected in 63 (70.8%) of the study hospitals. Most changes (41 out of 63) corresponded, as expected, to a switch from rhythm control to rate control.

The upper part of Table 4 shows the alternative estimates of the difference between 3-year mortality rates of patients treated with rhythm control drugs versus rate control drugs. The treatment effect is uniformly nonsignificant, but both point estimates and standard errors vary substantially. Among IV estimators, the change-time model yielded the point estimate closest to the null, the lowest standard error, and the most precise 95% confidence interval. The latter finding reflects the strongest correlation of the change-time-based instrument with the actual treatment (Table 4).

Finally, we performed a simple experiment by excluding 2 important measured confounders: age and secondary atrial fibrillation diagnosis at admission, both of which were associated with a higher mortality rate and with rate control treatment. The lower half of Table 4 shows a major impact of these exclusions on conventional estimates, which suggests a very significant reduction in mortality associated with rhythm control treatment. In contrast, the change-time IV model yielded a completely nonsignificant risk difference that was very similar to the original conventional estimate adjusted for the 2 confounders (model 1 in the upper half of Table 4) (i.e., it compensated for the bias artificially induced by these exclusions).

DISCUSSION

The proposed change-time model extends the physician prescribing-based IV originally developed by Brookhart et al. (13). In simulations, the change-time model reduced the variance of IV estimates, but its estimates were sometimes slightly biased. Similar bias due to modification of the original unbiased but numerically unstable estimator occurs, for example, in marginal structural models, where weight truncation reduces the variance but induces slight bias (36). Still, the RMSE comparisons indicated that our estimates were closer, on average, to the true treatment effect than were estimates based on either the last patient (model 2) (13) or all previous patients of the same physician (model 3). Indeed, in both simulations and the real-life atrial fibrillation example, the change-time model yielded the strongest instrument (i.e., had the highest correlation with the treatment

actually received by individual patients). Thus, the additional computations required by our method may be worth the effort.

In previous simulations, the inflated variance of the IV estimates often resulted in lower overall accuracy compared with conventional (biased) estimates (9). Yet, our change-time IV model substantially improved the variance-bias trade-off and, in simulations with 50–200 patients per physician, yielded substantially lower RMSEs than did the conventional model. Increasing the size of the physicians' practices increased both the sensitivity of detecting the preference changes and the instrument's strength. As expected, the gains in the accuracy offered by the change-time model over model 3 (all previous patients) increased with increasing frequency of preference changes. In the case of no or rare changes, our estimator should perform similarly to model 3.

Although the proposed change-time IV estimator produced, on average, the strongest instrument, the relative performance of different instruments varied across individual simulated samples. Furthermore, in empirical studies, the relative strength of alternative prescribing-preference-based IVs may vary depending on the frequency and pattern of changes over time in these preferences. Indeed, in some empirical studies, the instrument based on the last patient's prescription only was stronger than instruments based on larger numbers of previous patients (10, 15). For all of these reasons, the choice of the instrument most appropriate for a given application should be considered an empirical question. Researchers should apply several alternative IVs, including the instrument originally proposed by Brookhart et al. (13), its various modifications (10, 15, 32), and the change-time model we propose, and then compare their correlation with the actual treatment and the resulting treatment effect estimates.

The validity of both the original last patient-based IV (13) and our change-time-based IV depends on the validity of the assumptions underlying the IV method. Hernán and Robins (8) discussed these assumptions and their implications for the pharmacoepidemiologic applications of prescribing preference-based instruments, whereas Brookhart and Schneeweiss (37) present some evidence regarding their validity in their specific application. In general, our change-time model IV estimator will be valid whenever alternative IV estimators (10, 13, 14, 37, 38) are also valid.

Our method relies on somewhat arbitrary criteria to detect possible changes in physicians' preferences and estimate their times. Still, our model yielded RMSEs almost identical to those of the unrealistic model 4, which was based on the true occurrence and timing of change for each physician. As in all simulation studies, we investigated only a limited range of scenarios and relevant parameter values (39). Consistent with some empirical findings (13, 14, 32), we assumed strong prescribing preferences. Weaker preferences would increase the variance of all IV estimates and reduce their accuracy relative to conventional estimates but should not markedly alter the relative accuracy of alternative IV estimates. Indeed, the consistency of results across the simulated scenarios suggests that our conclusions are rather robust.

Both our change-time model and our simulated scenarios 1–12 assume an abrupt switch from a strong preference for one drug to another drug at a specific point in time. Yet, in practice, the prescribing preferences may often evolve more gradually. Still, our change-time estimator outperformed the 2 alternative IVs in both the simulated scenario 13, which assumed gradual, relatively slow, changes in preferences, and the real-life atrial fibrillation example. Thus, although it likely oversimplifies the actual, unknown pattern of changes, our model may still often improve the accuracy of IV estimates. In further research, investigators should develop and validate methods for estimating smooth changes in individual physicians' preferences over time. Similar to what was seen in the work of Cole and Hernán (36), arbitrary smooth patterns of time-dependent changes in probability of treatment could be modeled by spline functions (40).

There are no panaceas for addressing the potentially fatal impact of unobserved confounding on observational studies. Researchers should use different methods and link their possibly discrepant results with the underlying assumptions. The proposed change-time model may add a useful method to the IV toolbox and help to increase the accuracy of the estimates and conclusions of pharmacoepidemiologic studies.

REFERENCES

1. Abrahamowicz M, Tamblyn R. Drug utilization patterns. In: Armitage P, Colton T, eds. *Encyclopedia of Biostatistics*. Vol 3. 2nd ed. Chichester, United Kingdom: John Wiley & Sons Ltd; 2005:1533–1553.
2. Skegg DC. Evaluating the safety of medicines, with particular reference to contraception. *Stat Med*. 2001;20(23):3557–3569.
3. Brookhart MA, Stürmer T, Glynn RJ, et al. Confounding control in healthcare database research: challenges and potential approaches. *Med Care*. 2010;48(suppl 6):S114–S120.

4. Motheral BR, Fairman KA. The use of claims databases for outcomes research: rationale, challenges, and strategies. *Clin Ther.* 1997;19(2):346–366.

5. Walker AM. Confounding by indication. *Epidemiology.* 1996; 7(4):335–336.

6. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *J Am Stat Assoc.* 1996; 91(434):444–472.

7. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol.* 2000;29(4):722–729.

8. Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology.* 2006;17(4):360–372.

9. Ionescu-Ittu R, Delaney JA, Abrahamowicz M. Bias-variance trade-off in pharmacoepidemiological studies using physician-preference-based instrumental variables: a simulation study. *Pharmacoepidemiol Drug Saf.* 2009;18(7):562–571.

10. Rassen JA, Brookhart MA, Glynn RJ, et al. Instrumental variables II: instrumental variable application-in 25 variations, the physician prescribing preference generally was strong and reduced covariate imbalance. *J Clin Epidemiol.* 2009;62(12): 1233–1241.

11. Rassen JA, Schneeweiss S, Glynn RJ, et al. Instrumental variable analysis for estimation of treatment effects with dichotomous outcomes. *Am J Epidemiol.* 2009;169(3):273–284.

12. McMahon AD. Approaches to combat with confounding by indication in observational studies of intended drug effects. *Pharmacoepidemiol Drug Saf.* 2003;12(7):551–558.

13. Brookhart MA, Wang PS, Solomon DH, et al. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. *Epidemiology.* 2006; 17(3):268–275.

14. Schneeweiss S, Seeger JD, Landon J, et al. Aprotinin during coronary-artery bypass grafting and risk of death. *N Engl J Med.* 2008;358(8):771–783.

15. Hennessy S, Leonard CE, Palumbo CM, et al. Instantaneous preference was a stronger instrumental variable than 3- and 6-month prescribing preference for NSAIDs. *J Clin Epidemiol.* 2008;61(12):1285–1288.

16. Stafford RS, Furberg CD, Finkelstein SN, et al. Impact of clinical trial results on national trends in alpha-blocker prescribing, 1996–2002. *JAMA.* 2004;291(1):54–62.

17. Jackevicius CA, Anderson GM, Leiter L, et al. Use of the statins in patients after acute myocardial infarction: does evidence change practice? *Arch Intern Med.* 2001;161(2):183–188.

18. Wazana A. Physicians and the pharmaceutical industry: is a gift ever just a gift? *JAMA.* 2000;283(3):373–380.

19. Brookhart MA, Rassen JA, Schneeweiss S. Instrumental variable methods in comparative safety and effectiveness research. *Pharmacoepidemiol Drug Saf.* 2010;19(6):537–554.

20. Peay MY, Peay ER. The role of commercial sources in the adoption of a new drug. *Soc Sci Med.* 1988;26(12):1183–1189.

21. Søndergaard J, Vach K, Kragstrup J, et al. Impact of pharmaceutical representative visits on GPs' drug preferences. *Fam Pract.* 2009;26(3):204–209.

22. Dormuth CR, Maclure M, Bassett K, et al. Effect of periodic letters on evidence-based drug therapy on prescribing behaviour: a randomized trial. *CMAJ.* 2004;171(9):1057–1061.

23. Gelman A, Hill J. Causal inference using more advanced methods. In: Gelman A, Hill J, eds. *Data Analysis Using Regression and Multilevel/Hierarchical Models.* New York, NY: Cambridge University Press; 2007:215–225.

24. Staiger D, Stock JH. Instrumental variables regression with weak instruments. *Econometrica.* 1997;65(3):557–586.

25. Pawitan Y. Change-point problem. In: Armitage P, Colton T, eds. *Encyclopedia of Biostatistics.* Hoboken, NJ: John Wiley & Sons, Inc; 2005. (doi: 10.1002/0470011815.b2a12011).

26. Akaike H. A new look at the statistical model identification. *IEEE T Automatic Contr.* 1974;AC-19(6):716–723.

27. Tamblyn R, Lavoie G, Petrella L, et al. The use of prescription claims databases in pharmacoepidemiological research: the accuracy and comprehensiveness of the prescription claims database in Québec. *J Clin Epidemiol.* 1995;48(8):999–1009.

28. Wyse DG, Waldo AL, DiMarco JP, et al. A comparison of rate control and rhythm control in patients with atrial fibrillation. *N Engl J Med.* 2002;347(23):1825–1833.

29. Benjamin EJ, Wolf PA, D'Agostino RB, et al. Impact of atrial fibrillation on the risk of death: the Framingham Heart Study. *Circulation.* 1998;98(10):946–952.

30. Wolf PA, D'Agostino RB, Belanger AJ, et al. Probability of stroke: a risk profile from the Framingham Study. *Stroke.* 1991;22(3):312–318.

31. Rassen JA, Mittleman MA, Glynn RJ, et al. Safety and effectiveness of bivalirudin in routine care of patients undergoing percutaneous coronary intervention. *Eur Heart J.* 2010;31(5): 561–572.

32. Pratt N, Roughead EE, Ryan P, et al. Antipsychotics and the risk of death in the elderly: an instrumental variable analysis using two preference based instruments. *Pharmacoepidemiol Drug Saf.* 2010;19(7):699–707.

33. Havranek EP, Wolfe P, Masoudi FA, et al. Provider and hospital characteristics associated with geographic variation in the evaluation and management of elderly patients with heart failure. *Arch Intern Med.* 2004;164(11):1186–1191.

34. O'Connor GT, Quinton HB, Traven ND, et al. Geographic variation in the treatment of acute myocardial infarction: the Cooperative Cardiovascular Project. *JAMA.* 1999;281(7): 627–633.

35. Nair H, Manchanda P, Bhatia T. *Asymmetric Peer Effects in Physician Prescription Behavior: The Role of Opinion Leaders.* (Research Paper Series, no. 1970). Stanford, CA: Stanford Graduate School of Business; 2006.

36. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol.* 2008;168(6):656–664.

37. Brookhart MA, Schneeweiss S. Preference-based instrumental variable methods for the estimation of treatment effects: assessing validity and interpreting results. *Int J Biostat.* 2007; 3(1):14.

38. Schneeweiss S, Setoguchi S, Brookhart A, et al. Risk of death associated with the use of conventional versus atypical antipsychotic drugs among elderly patients. *CMAJ.* 2007;176(5): 627–632.

39. Burton A, Altman DG, Royston P, et al. The design of simulation studies in medical statistics. *Stat Med.* 2006;25(24): 4279–4292.

40. Abrahamowicz M, MacKenzie TA. Joint estimation of time-dependent and non-linear effects of continuous covariates on survival. *Stat Med.* 2007;26(2):392–408.