Group-based Trajectory Models

A New Approach to Classifying and Predicting Long-Term Medication Adherence

Jessica M. Franklin, PhD, William H. Shrank, MD,* Juliana Pakes, MEd,* Gabriel Sanfélix-Gimeno, PhD, PharmD,*†‡ Olga S. Matlin, PhD,§ Troyen A. Brennan, MD, JD,§ and Niteesh K. Choudhry, MD, PhD**

Background: Classifying medication adherence is important for efficiently targeting adherence improvement interventions. The purpose of this study was to evaluate the use of a novel method, group-based trajectory models, for classifying patients by their long-term adherence.

Research Design: We identified patients who initiated a statin between June 1, 2006 and May 30, 2007 in prescription claims from CVS Caremark and evaluated adherence over the subsequent 15 months. We compared several adherence summary measures, including proportion of days covered (PDC) and trajectory models with 2–6 groups, with the observed adherence pattern, defined by monthly indicators of full adherence (defined as having ≥ 24 d covered of 30). We also compared the accuracy of adherence prediction based on patient characteristics when adherence was defined by either a trajectory model or PDC.

Results: In 264,789 statin initiators, the 6-group trajectory model summarized long-term adherence best ($C = 0.938$), whereas PDC summarized less well ($C = 0.881$). The accuracy of adherence predictions was similar whether adherence was classified by PDC or by trajectory model.

Conclusions: Trajectory models summarized adherence patterns better than traditional approaches and were similarly predicted by covariates. Group-based trajectory models may facilitate targeting of interventions and may be useful to adjust for confounding by health-seeking behavior.

Key Words: adherence, comparative effectiveness, latent class, longitudinal data

(*Med Care* 2013;51: 789–796)

Medication adherence, or the degree to which patients take medications as prescribed by their health care providers, has been shown to be poor across a range of conditions and medications.[1–5] Nonadherence is associated with adverse health outcomes and higher health care costs.[1] Classifying patients by their adherence to a specific medication can increase the efficiency of interventions to improve adherence by facilitating the targeting of interventions to patients most likely to benefit.[6] In addition, grouping patients according to their medication adherence can aid in evaluating treatment efficacy in the presence of nonadherence or in understanding heterogeneity in treatment effectiveness.

In administrative claims data, adherence is most often assessed through measures such as the proportion of days covered (PDC), defined as the number of follow-up days covered with medication divided by the total number of days in follow-up.[7–9] The advantage of PDC or similar measures is that they reduce a potentially complex pattern of longitudinal adherence observations down to a single number. However, as a result of this dimension reduction, they are limited in their ability to distinguish between different adherence experiences. For example, identical PDC values may be calculated for patients with (1) consistent use early during follow-up but poor subsequent use, (2) poor initial use and consistent subsequent use, or (3) intermittent adherence throughout follow-up. These differences are poorly captured by possession measures but can have important implications for patient prognosis, particularly when adherence is assessed over longer time periods. Similarly, measures of medication persistence (eg, the number of months of continuous medication use) generally cannot differentiate between patients that stop medication completely versus patients that simply have a gap in use.

Group-based trajectory models[10,11] may provide an alternative method for summarizing long-term medication

From the *Department of Medicine, Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women's Hospital, Harvard Medical School; †Department of Epidemiology, Harvard School of Public Health, Boston, MA; ‡Centro Superior de Investigación en Salud Pública, Valencia, Spain; and §CVS Caremark, Woonsocket, RI.

Supported by an unrestricted grant from CVS Caremark to the Brigham and Women's Hospital. G.S-G. was partially funded by the ASISA Harvard Fellowship for Excellence in Clinical Research and the Carlos III Health Institute (BA11/00053). O.S.M. and T.A.B. are employees of CVS Caremark.

The authors declare no conflict of interest.

Reprints: Jessica M. Franklin, PhD, Department of Medicine, Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women's Hospital, Harvard Medical School, 1620 Tremont St, Suite 3030, Boston, MA 02120. E-mail: jmfranklin@partners.org.

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Website, www.lww-medicalcare.com.

ISSN: 0025-7079/13/5109-0789

adherence accounting for the dynamic nature of adherence over time. These models estimate the change over time in an outcome that is measured repeatedly, and they are designed to identify individuals who have similar longitudinal response patterns. Trajectory models and related methods for identifying developmental trajectories have been used often in sociological and medical research, including research on health-related behaviors,[12–14] childhood development,[15–17] and disease progression,[18–20] but have had limited application in the study of prescription medications.[21,22] The objective of this study was to evaluate the use of group-based trajectory models for classifying patients by their long-term medication adherence using a large pharmacy claims database.

METHODS

Cohort

We identified patients who initiated a statin medication between June 1, 2006 and May 30, 2007 in prescription claims from CVS Caremark (Woonsocket, RI), a pharmacy benefit manager.[23] We defined the index date as the date of the first prescription fill for any statin in the study period. Patients were required to have continuous pharmacy benefits coverage for 1 year before the index date to ensure no prior statin use during this period. Patients were also required to have 15 months of enrollment after the index date to allow complete adherence follow-up. We excluded patients who received any statin prescription through mail, as these patients may have systematically different adherence experiences from patients that fill their statin prescriptions only at retail pharmacies. The institutional review board of Brigham and Women's Hospital approved the study.

Measuring Adherence

We followed patients and recorded fills for prescriptions of any statin for 15 months after the index date. For each patient, we created a "supply diary" that indicates whether each day was covered with medication by linking all observed statin fills based on the dispensing date and the days' supply. When a patient was dispensed a statin before the end of the days' supply for the previous dispensing, we assumed that use of the new statin refill began the day after the days' supply for the prior statin fill had elapsed. If a patient accumulated >180 days' supply on a given day, we truncated the accumulated supply at 180 days.

On the basis of the supply diary, we created 15 binary indicator variables that indicate for each month of adherence follow-up whether that month had at least 24 days (ie, 80% of each 30 d period) covered. We focus on 80% coverage in each month because this coverage is believed to be sufficient in the short term to achieve the clinical benefits of statins,[24–26] and quality measures (Medicare STAR ratings, National Committee for Quality Assurance, National Quality Forum) define adherence in this way. These binary indicators defined the observed adherence pattern during follow-up. Also based on the supply diary, we calculated PDC over the 15 months of follow-up, as well as the proportion of months covered (PMC), defined as the number of months with at least 24 days covered divided by the total number of months of follow-up (15). We did not evaluate persistence measures, as these measures by definition ignore medication use after discontinuation and therefore would not be expected to accurately summarize use during that period.

Group-based Trajectory Models

We used group-based trajectory models to classify patients by their observed medication adherence. We focused on this method because it is simple to implement using "Proc Traj," a free downloadable add-on package to base SAS (SAS, version 9.2, Cary, NC), and because it was shown to be superior for identifying underlying longitudinal trajectories.[22] In a trajectory model, several regression models are estimated simultaneously through maximization of a likelihood that combines the information from all models. Specifically, based on individuals' adherence patterns over time, the probability of belonging to each potential adherence group is modeled as a simple multinomial logistic regression with no predictors (only an intercept for each group). Within each adherence group, adherence is modeled as a smooth function of time using up to a fourth order polynomial. Therefore, the model assumes that repeated observations on the same individual are independent conditional on trajectory group, meaning that the within-person correlation structure is explained completely by the estimated trajectory curve for each person's group. The output of a group-based trajectory model includes estimated probabilities of group membership for each individual and each group and an estimated trajectory curve over time for each group.

We modeled the 15 monthly binary indicators of adherence as the longitudinal outcome, and we used logistic regression for the group-specific models. The time variable was months since the index date (1–15). We estimated models using between 2 and 6 groups under the assumption that interpreting or utilizing >6 adherence groups would be prohibitively difficult. In each model, we allowed all groups to have "order 3," indicating the use of cubic terms of time to model the probability of being adherent in each group. We considered the use of other model orders (2 or 4), but found that order 3 provided the best fit (based on comparisons of estimated and observed adherence trajectories). On the basis of these models, we estimated the probability of membership in each group for each individual, and we assigned patients to the trajectory group with the highest membership probability.

Covariates

On the basis of prior adherence research,[1,23,27–31] we assessed several patient characteristics as of the index date for predicting subsequent adherence to statins. We focused on covariates measured on or before the initiation of statins to assess the prediction accuracy that was possible before observing any follow-up and to avoid including as predictors changes in patient health status that may themselves be consequences of the use (or nonuse) of the medication under study. We included all measured covariates in the prediction models and did not perform variable selection, as (1) we had a very large study size, (2) there were relatively few covariates available, and (3) variable selection is generally not necessary when the goal of modeling is prediction as

opposed to inference.[32] All analyses were restricted to patients with complete covariate information.

From enrollment files, we ascertained patient age, sex, region of residence in the United States, and type of prescription benefit coverage, classified as employer sponsored, health insurer carve out (ie, beneficiaries who are fully insured through a commercial health insurer but whose prescription drug coverage is "carved out" and provided separately by a pharmacy benefit manager), Medicare, or other (including Medicaid beneficiaries, cash card holders, and offshore customers). By linking to US Census data, we assessed the median income, the proportion of black residents, and the proportion of high school graduates in the zip code of residence for each patient. We recorded features of the index statin prescription, including whether it was for >30 days, the patient copayment, and the sum of copayments for any other medications filled on the index date. Finally, in the 6 months before index we assessed the total copayments for all medications, the number of distinct drugs (defined by GPI-8 codes), and prescription fills of drugs that indicate the presence of selected comorbidities, including congestive heart failure, hypertension, stroke, chronic obstructive pulmonary disease or asthma, depression, and diabetes mellitus.

Comparing Adherence Summaries

We assigned an adherence summary value to each person-month in the data using each adherence measurement method. For PDC, PDC dichotomized at $\geq/<80\%$, and PMC, the summaries vary across individuals but are constant over time. For trajectory models, the summary values were taken from the estimated probability of adherence over time within each group, so the summary values vary across time but are constant across individuals assigned to the same group. We then calculated the mean, median, and interquartile range of each adherence summary, separately in person-months that were fully covered or not fully covered with a statin. We calculated *C*-statistics to measure the correspondence between the adherence pattern (the "gold standard"), defined by the 15 monthly indicators, and each adherence summary (the "test" values). A *C*-statistic of 1.0 indicates perfect correspondence, whereas a value of 0.5 indicates no association. To choose among the estimated trajectory models, we compared the Bayesian Information Criterion (BIC), as recommended in the trajectory model literature,[33] and the estimated probabilities of group membership with lower BIC values and higher probabilities indicating better model fit.

Predicting Adherence Classes

To compare the accuracy of adherence prediction across the varying methods for grouping adherence, we created several binarizations of adherence. In each trajectory model, we binarized trajectory groups as best adherence (group 1) versus all other groups. In the 6-group model, we additionally created binary groupings to evaluate the prediction of each unique group (eg, group 2 vs. all others, group 3 vs. all others, etc.). We also created several binary groupings based on PDC, including PDC dichotomized at 80% and each sextile of PDC versus all others (in order to be comparable to the 6-group trajectory model). For each binarization, we estimated a multivariable logistic regression model with the binary adherence indicator as the outcome and all patient characteristics as predictors.

RESULTS

Adherence Summaries

In 264,789 statin initiators, we observed 14,832 unique adherence patterns of $2^{15}=32{,}768$ that were possible. On the basis of these patterns, the group-based trajectory models identified distinct patterns of medication adherence (Figs. 1, 2). When allowing for 6 groups (Fig. 2), the trajectory model identified individuals that (1) were nearly always adherent (23.4% of the study population); (2) had a brief gap in medication use or filled irregularly during the first 9 months, but improved during the last 6 months (11.4%); (3) had slowly declining adherence throughout the 15 months (11.3%); (4) used statins only occasionally across the 15 months (15.0%); (5) had a rapid decline in statin use after initiation (19.3%); and (6) had virtually no fills after their index prescription fill (23.4%).

In Figures 1 and 2, separately for each trajectory model, we also present the median and interquartile range of the predicted probability of membership in each trajectory group among individuals assigned to that group. For example, the values at the top of each plot describe the predicted probability of being in trajectory 1, among those individuals that were assigned to trajectory 1. As group assignments are based on the probabilities of group membership, these probabilities should be relatively high, but may still vary considerably. For example, in a model with 5 groups, an individual could hypothetically be assigned to trajectory 1 based on a probability as low as 0.21, as long as the probabilities of other groups are lower (ie, probabilities of 0.2, 0.2, 0.2, and 0.19 for groups 2–5, respectively). As shown in the figures, trajectory group membership is predicted with high probability in all models, although the predicted probabilities are generally more variable as the number of groups increases. Patients assigned to trajectories representing either consistently good or consistently poor adherence tended to have a very high probability of membership in their respective trajectory group.

Table 1 compares the trajectory models and other adherence summaries. In the approximately 2.1 million person-months that were not fully covered, better-performing adherence summaries would have low values (near 0). Of the adherence summaries considered, PDC dichotomized at 80% had the lowest mean value (0.05), whereas continuous PDC had the highest (0.31). All trajectory models had mean summary values between 0.19 and 0.25. In the nearly 1.7 million person-months that were fully covered, better-performing adherence summaries would have high values (near 1). In this group, the 6-group trajectory model had the highest mean value (0.79), and dichotomized PDC had the lowest (0.58).

Overall, the 6-group trajectory model summarized long-term adherence best ($C=0.938$), whereas dichotomized PDC summarized worst ($C=0.762$). PMC summarized observed adherence slightly better than PDC ($C=0.887$ and

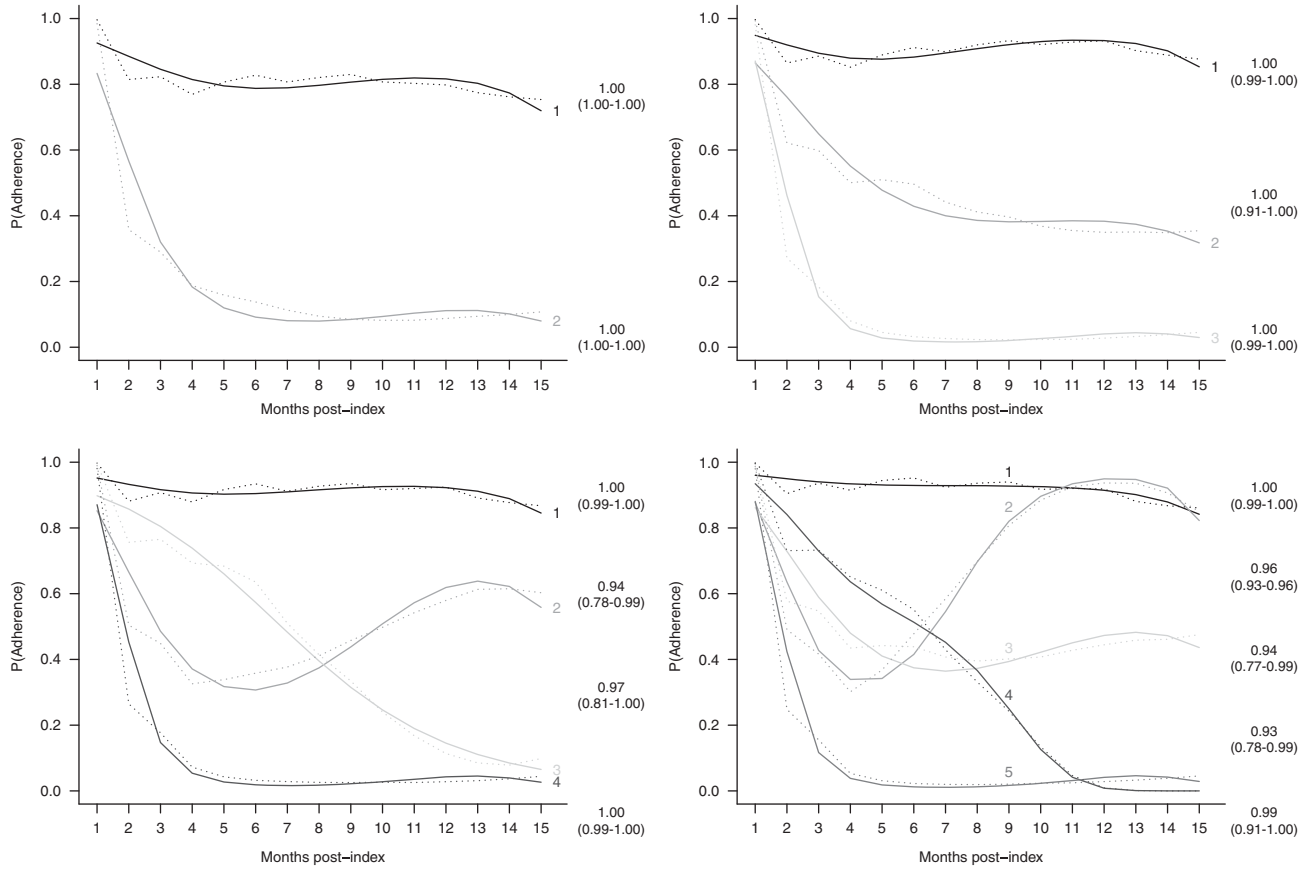

FIGURE 1. Trajectory models using 2–5 groups. In each plot, the predicted probability of adherence in each group is plotted with solid lines. The observed proportion of individuals in each group that are adherent is plotted with dotted lines. Median (and interquartile range) trajectory membership probabilities by group assignment are displayed at the right of each plot.

0.881, respectively), which was expected as PMC was based on the monthly indicators of having $\geq 80\%$ of days covered. Among the 5 trajectory models, the *C*-statistic increased with increasing number of groups, but even the 2-group model ($C = 0.873$) outperformed the dichotomized PDC measure. The 6-group trajectory model also had the lowest BIC value, indicating that there was no evidence of over fitting. Because of our very large study size, BIC would likely continue to decrease in models with >6 groups, but ≥ 7 adherence groups would be difficult to interpret clinically.

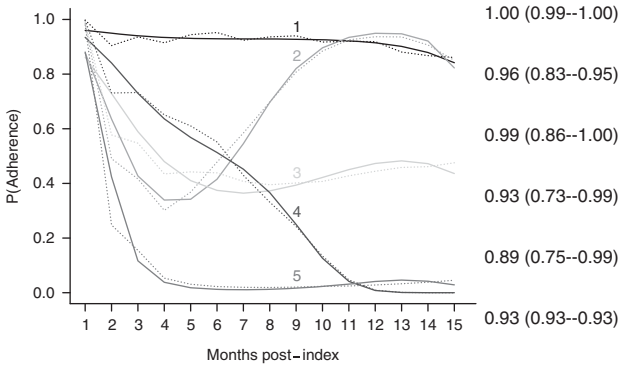

FIGURE 2. Trajectory model using 6 groups. The predicted probability of adherence in each group is plotted with solid lines. The observed proportion of individuals in each group that are adherent is plotted with dotted lines. Median (and interquartile range) trajectory membership probabilities by group assignment are displayed at the right.

Predicting Adherence Classes

The characteristics of patients in each trajectory from the 6-group model are shown in Table 2. Patients with the best adherence (trajectory 1) were on average older, more likely to have a higher income, more likely to be a high school graduate and less likely to be black than patients in other trajectory groups. In addition, these patients were more likely to be Medicare beneficiaries, to live in New England, and to have nearly all of the predefined comorbidities. Patients with the lowest adherence (trajectory 6) were generally younger, more likely to be male, and much less likely to have >30 days' supply in their index statin fill compared with patients in other trajectory groups.

The *C*-statistics from logistic regression models using these patient characteristics to predict adherence are presented in Table 3. Given the limited number and scope of patient characteristics available in this study, prediction accuracy based on the baseline covariates was poor for all adherence groupings.

TABLE 1. Comparisons of Adherence Summaries With Observed Adherence Pattern

Adherence Measure	Summary Values*: Not Fully Adherent N = 2,105,593	Summary Values*: Fully Adherent N = 1,866,242	*C*-statistic†	BIC‡
PDC	0.31/0.23 (0.13–0.46)	0.76/0.85 (0.61–0.95)	0.881	—
PDC ≥ 80%	0.05/0.00 (0.00–0.00)	0.58/1.00 (0.00–1.00)	0.762	—
PMC	0.25/0.20 (0.07–0.33)	0.72/0.80 (0.53–0.93)	0.887	—
Trajectory models				
2 groups	0.25/0.10 (0.08–0.32)	0.73/0.80 (0.79–0.82)	0.873	2024356
3 groups	0.22/0.04 (0.03–0.38)	0.75/0.88 (0.55–0.92)	0.908	1902300
4 groups	0.20/0.05 (0.03–0.33)	0.77/0.90 (0.64–0.92)	0.921	1862769
5 groups	0.20/0.04 (0.02–0.41)	0.77/0.90 (0.59–0.93)	0.924	1835308
6 groups	0.19/0.05 (0.00–0.31)	0.79/0.92 (0.72–0.95)	0.938	1818452

*Mean/median (interquartile range) of the value of each adherence summary as assigned to fully adherent person-months (at least 24 d covered) and not fully adherent person-months.

†*C*-statistics compare the adherence summaries (the "test" values) to the observed adherence (the "gold standard") in each person-month. Higher values (near 1.0) indicate summaries that better discriminate between adherent and nonadherent months. Low values (near 0.5) indicate little discriminatory ability.

‡This value is based on the model likelihood with a penalty for the number of model parameters. It is not directly interpretable, but lower values indicate better model fit.

BIC indicates Bayesian information criterion; PDC, proportion of days covered; PMC, proportion of months covered.

The highest *C*-statistic from all models was 0.665, which came from the model that predicted trajectory 6 (poor adherence) versus all others. Predicting poor adherence based on PDC (sextile 6 vs. all others) had similar results ($C = 0.656$). Among the models that predicted good versus bad adherence, the highest prediction accuracy came from the model with the 2-group trajectory model groupings as the outcome ($C = 0.649$), but the *C*-statistics were similar for all others. Prediction accuracy was lower when attempting to predict intermediate adherence groups, based on either trajectory models or PDC.

DISCUSSION

In this large cohort of statin initiators, we found that group-based trajectory models summarized the observed 15 months of longitudinal adherence to statins more accurately than the conventional PDC approach. In particular, even when using only a 2-group model, the adherence summary was similar in accuracy to the continuous PDC measure and was significantly better than dichotomized PDC. This finding indicates that the adherence groups identified by the trajectory models are generally more homogenous with respect to observed adherence than groups based on PDC. In addition, when using baseline covariates to predict adherence classifications, we found that prediction accuracy was similar whether the adherence classification was based on a trajectory model or on PDC. Therefore, the association between covariates and adherence is not diminished with the use of trajectory models.

When comparing PDC and trajectory models for summarizing medication adherence, we used multiple models with varying numbers of groups. As recommended in the trajectory model literature,[10,11,22,34,35] we used BIC and estimated group membership probabilities to choose among the models and to focus on the 6-group model for further evaluations. However, the appropriate number of groups in a given application may partly depend on how the adherence groups are to be used. In some cases, a simple dichotomization of acceptable versus unacceptable adherence may be preferred, so that a 2-group model is sufficient; in other cases, a more refined classification may be preferred. We also used third order polynomials for modeling adherence over time in all models. This choice is largely dependent on the number of repeated observations; we have found that with 12–15 months of data, order 3 is sufficiently flexible to capture meaningful changes in medication use over time. With only 6 or 9 months, third order polynomials may overfit the available data, and second order models may be preferred.

In this paper, we have focused on group-based latent trajectory models[34] for identifying groups based on longitudinal observations, but other methods with the same goal are available, as discussed by Twisk and Hoekstra.[22] The models that we present are equivalent to the latent class growth analysis models discussed in their paper, which performed favorably over other methods considered, including latent class growth mixture modeling,[36] K-means clustering,[37] and latent class analysis.[38] The approach that we used has other favorable properties as well, as it has a simple, freely available macroimplementation for SAS, and it can accommodate missing adherence data without additional analysis steps.

In our study, adherence was observed completely for all patients because we restricted the cohort to patients with 15 months of follow-up. If we had included patients with less follow-up, they would have had missing adherence data for months remaining after the end of follow-up. These patients can still be used for model estimation, and they would also be assigned to trajectory groups, but generally with lower probabilities than the group assignments for patients with complete data. Thus, the high predicted probabilities of group membership that were observed in our example could partly be attributed to the fact that we had complete adherence data on all patients.

We also restricted the study to patients with complete covariate information. Only a small proportion of the original cohort was lost due to missing covariate information (8%). These patients were generally very similar to patients with complete covariate information with respect to observed

TABLE 2. Patient Characteristics by Trajectory Using a 6-Group Model

Characteristics	Trajectory 1 N = 61,962	Trajectory 2 N = 30,055	Trajectory 3 N = 29,994	Trajectory 4 N = 39,835	Trajectory 5 N = 51,226	Trajectory 6 N = 51,717
Age [mean (SD)] (y)	61.1 (13.1)	59.3 (13.2)	58.7 (13.5)	56.7 (13.1)	57.0 (13.8)	55.6 (13.8)
Male sex	48.7	48.6	49.2	49.7	49.8	51.3
Income mean (SD)	54,750 (20,237)	53,605 (19,841)	53,099 (19,456)	52,224 (19,278)	52,046 (18,808)	51,105 (18,698)
Percent black [mean (SD)]	9.7 (16.9)	11.7 (19.1)	11.7 (19.2)	14.4 (21.9)	15.4 (22.7)	13.4 (20.9)
Percent HSG [mean (SD)]	81.8 (16.0)	80.8 (11.0)	80.6 (11.0)	79.6 (11.6)	78.6 (12.0)	79.5 (11.6)
Insurance status						
Employer sponsored	49.4	51.0	49.7	53.3	52.2	54.1
Health plan	18.8	19.6	21.2	22.0	21.0	21.2
Medicare	27.8	24.8	23.2	19.2	18.7	15.5
Other	4.0	4.6	6.0	5.5	8.1	9.2
Geographic region						
West/Southwest	23.0	23.3	23.6	24.2	23.9	25.2
Southeast	27.1	30.1	30.6	32.2	32.4	33.1
New England	10.3	8.9	8.1	7.8	6.8	6.0
Mid Atlantic	12.7	12.6	12.6	12.1	12.9	12.7
Mid West	27.0	26.1	25.1	23.7	23.9	23.0
Comorbidities						
Congestive heart failure	48.4	44.0	42.2	39.3	37.4	33.7
Hypertension	54.1	50.1	47.8	45.5	43.2	39.1
Stroke	7.5	6.7	6.6	5.1	5.3	4.5
COPD/asthma	9.6	9.8	9.6	8.9	9.2	8.8
Depression	19.8	19.6	20.4	17.3	17.6	15.1
Diabetes	18.6	18.8	18.1	19.0	16.3	15.9
GPI-8 [mean (SD)]	5.8 (4.5)	5.6 (4.3)	5.6 (4.4)	5.3 (4.2)	5.3 (4.1)	5.1 (4.0)
Index days supply >30	14.7	14.4	16.7	9.4	14.6	1.1
Copay baseline ($)						
0	17.4	17.9	18.3	18.3	19.4	20.4
0.1–19.99	9.0	9.7	9.5	11.3	11.0	12.0
20–99.99	26.6	27.9	27.3	29.1	28.8	29.5
100–199.99	16.1	16.5	16.3	16.7	15.7	15.9
≥ 200	30.9	27.9	28.7	24.7	25.2	22.2
Copay index statin ($)						
0	7.6	6.9	7.1	6.1	6.9	7.0
0.1–9.99	33.7	31.1	30.1	28.0	27.2	25.8
10–24.99	34.1	34.5	34.4	35.9	34.7	35.0
≥ 25	24.7	27.4	28.3	30.1	31.2	32.2
Copay index nonstatin ($)						
0	53.2	53.6	52.6	52.9	53.7	52.7
0.1–19.99	24.0	23.3	22.5	23.3	22.7	23.3
≥ 20	23.7	24.2	25.0	23.8	23.6	24.1

COPD indicates chronic obstructive pulmonary disease; HSG, high school graduate.

baseline characteristics, but they did have slightly lower adherence (mean PDC 47.5% vs. 48.7%). The estimated trajectories were very similar when including all patients (see supplementary material, Supplementary Digital Content 1, http://links.lww.com/MLR/A493), but the distribution of patients across trajectories was slightly different. Therefore, the estimates of prediction accuracy may have been different if we had included patients with missing covariate data.

One limitation of trajectory models as applied to pharmacy claims data is that they may misclassify patients that fill prescriptions without actually taking them. This problem is particularly acute at the beginning of follow-up, where filling a single 90-day prescription guarantees that we observe full adherence for the first 3 months, even if the patient never takes the medication. However, this issue is common to all methods of measuring adherence based on pharmacy claims and is not unique to trajectory models. Furthermore, this issue becomes less important as the period of adherence follow-up is lengthened.

If identifying groups of patients with similar adherence is not of primary importance, another option is to analyze the longitudinal adherence patterns directly with repeated measures techniques. However, this approach is often difficult in practice. If the study goal is to predict adherence, then models must predict each measure of adherence, which may be noninformative if only baseline covariates are available. If the goal is to use observed adherence to predict future clinical outcomes, then investigators must interpret a separate effect of adherence in each month. In contrast, reducing the longitudinal adherence pattern into a single number or grouping reduces model complexity and improves interpretability of conclusions. PDC has been the standard metric for achieving this simplification, but trajectory models provide a flexible alternative.

On the basis of the findings in this study, group-based trajectory models may facilitate research on medication adherence and medication effectiveness. Trajectory models can identify patients with distinct patterns of medication fills so

TABLE 3. *C*-Statistics Measuring the Ability of Each Potential Adherence Binarization to be Predicted by Baseline Covariates

Classification	Based on	Binarized as	*C*-statistic
Good	PDC	>80%	0.642
Good	2-group trajectory	1 vs. 2	0.649
Good	3-group trajectory	1 vs. 2–3	0.646
Good	4-group trajectory	1 vs. 2–4	0.647
Good	5-group trajectory	1 vs. 2–5	0.646
Good	6-group trajectory	1 vs. others	0.644
Intermediate	6-group trajectory	2 vs. others	0.562
Intermediate	6-group trajectory	3 vs. others	0.567
Intermediate	6-group trajectory	4 vs. others	0.563
Intermediate	6-group trajectory	5 vs. others	0.570
Poor	6-group trajectory	6 vs. others	0.665
Good	PDC sextile	1 vs. others	0.652
Intermediate	PDC sextile	2 vs. others	0.591
Intermediate	PDC sextile	3 vs. others	0.542
Intermediate	PDC sextile	4 vs. others	0.535
Intermediate	PDC sextile	5 vs. others	0.577
Poor	PDC sextile	6 vs. others	0.656

Both trajectory groups and PDC sextiles are labeled so that increasing group number is associated with lower adherence. The first column identifies what type of adherence classification was predicted (good adherence vs. all others, an intermediate adherence group vs. all others, or poor adherence vs. all others), the second column indicates which adherence summary was used as the basis of the dichotomous outcome (either PDC or a trajectory model), and the third column indicates which dichotomization was used.

PDC indicates proportion of days covered.

that investigators can target different interventions to patients with different adherence experiences. They may improve the prediction of adherence to a new medication by better summarizing the prior use of chronic medications. The groups identified by trajectory models may also be useful when accounting for patients' health-seeking behavior in comparative effectiveness studies, as the different adherence trajectories may represent unique strata of health consciousness. Finally, trajectory models could be used to improve understanding of the effects of nonadherence on subsequent clinical outcomes, as trajectory groupings incorporate both quantity and timing of medication availability. Future work should focus on investigating these associations.

REFERENCES

1. Choudhry NK, Setoguchi S, Levin R, et al. Trends in adherence to secondary prevention medications in elderly post-myocardial infarction patients. *Pharmacoepidemiol Drug Saf*. 2008;17:1189–1196.
2. Osterberg L, Blaschke T. Adherence to medication. *N Engl J Med*. 2005;353:487–497.
3. Brookhart MA, Patrick AR, Dormuth C, et al. Adherence to lipid-lowering therapy and the use of preventive health services: an investigation of the healthy user effect. *Am J Epidemiol*. 2007;166:348–354.
4. Caro JJ, Salas M, Speckman JL, et al. Persistence with treatment for hypertension in actual practice. *Can Med Assoc J*. 1999;160:31–37.
5. Solomon DH, Avorn J, Katz JN, et al. Compliance with osteoporosis medications. *Arch Intern Med*. 2005;165:2414–2419.
6. Cutrona SL, Choudhry NK, Fischer MA, et al. Targeting cardiovascular medication adherence interventions. *J Am Pharm Assoc*. 2012;52:381–397.
7. Ho PM, Bryson CL, Rumsfeld JS. Medication adherence. *Circulation*. 2009;119:3028–3035.
8. Steiner JF, Koepsell TD, Fihn SD, et al. A general method of compliance assessment using centralized pharmacy records: description and validation. *Med Care*. 1988;26:814–823.
9. Steiner JF, Prochazka AV. The assessment of refill compliance using pharmacy records: methods, validity, and applications. *J Clin Epidemiol*. 1997;50:105–116.
10. Nagin D. *Group-based Modeling of Development*. Cambridge, MA: Harvard University Press; 2005.
11. Nagin DS. Analyzing developmental trajectories: a semiparametric, group-based approach. *Psychol Methods*. 1999;4:139–157.
12. Caldeira KM, O'Grady KE, Garnier-Dykstra LM, et al. Cigarette smoking among college students: longitudinal trajectories and health outcomes. *Nicotine Tob Res*. 2012;14:777–785.
13. Conklin CA, Perkins KA, Sheidow AJ, et al. The return to smoking: 1-year relapse trajectories among female smokers. *Nicotine Tob Res*. 2005;7:533–540.
14. Hong SI, Hasche L, Bowland S. Structural relationships between social activities and longitudinal trajectories of depression among older adults. *The Gerontologist*. 2009;49:1–11.
15. Brame B, Nagin DS, Tremblay RE. Developmental trajectories of physical aggression from school entry to late adolescence. *J Child Psychol Psychiatry*. 2001;42:503–512.
16. Broidy LM, Nagin DS, Tremblay RE, et al. Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: a six-site, cross-national study. *Dev Psychol*. 2003;39:222–245.
17. Castellanos FX, Lee PP, Sharp W, et al. Developmental trajectories of brain volume abnormalities in children and adolescents with attention-deficit/hyperactivity disorder. *JAMA*. 2002;288:1740–1748.
18. Verkleij SPJ, Hoekstra T, Rozendaal RM, et al. Defining discriminative pain trajectories in hip osteoarthritis over a 2-year time period. *Ann Rheum Dis*. 2012;71:1517–1523.
19. Walsh K, Danielson CK, McCauley J, et al. Longitudinal trajectories of posttraumatic stress disorder symptoms and binge drinking among adolescent girls: the role of sexual victimization. *J Adolesc Health*. 2012;50:54–59.
20. Yeates KO, Taylor HG, Rusin J, et al. Longitudinal trajectories of postconcussive symptoms in children with mild traumatic brain injuries and their relationship to acute clinical status. *Pediatrics*. 2009;123:735–743.
21. Riegel B. Predictors of objectively measured medication nonadherence in adults with heart failure. *Circ Heart Fail*. 2011;5:430–436.
22. Twisk J, Hoekstra T. Classifying developmental trajectories over time should be done with great caution: a comparison between methods. *J Clin Epidemiol*. 2012;65:1078–1087.
23. Choudhry NK, Fischer MA, Avorn J, et al. The implications of therapeutic complexity on adherence to cardiovascular medications. *Arch Intern Med*. 2011;171:814–822.
24. Benner JS, Glynn RJ, Mogun H, et al. Long-term persistence in use of statin therapy in elderly patients. *JAMA*. 2002;288:455–461.
25. Karve S, Cleves MA, Helm M, et al. Good and poor adherence: optimal cut-point for adherence measures using administrative claims data. *Curr Med Res Opin*. 2009;25:2303–2310.
26. Shrank WH, Hoang T, Ettner SL, et al. The implications of choice: prescribing generic or preferred pharmaceuticals improves medication adherence for chronic conditions. *Arch Intern Med*. 2006;166:332–337.
27. Chan DC, Shrank WH, Cutler D, et al. Patient, physician, and payment predictors of statin adherence. *Med Care*. 2010;48:196–202.
28. Choudhry NK, Shrank WH, Levin RL, et al. Measuring concurrent adherence to multiple related medications. *Am J Manag Care*. 2009;15:457–464.
29. Fischer MA, Choudhry NK, Brill G, et al. Trouble getting started: predictors of primary medication nonadherence. *Am J Med*. 2011;124:1081.e9–1081.e22.
30. Kulik A, Shrank WH, Levin R, et al. Adherence to statin therapy in elderly patients after hospitalization for coronary revascularization. *Am J Cardiol*. 2011;107:1409–1414.
31. Setoguchi S, Choudhry N, Levin R, et al. Temporal trends in adherence to cardiovascular medications in elderly patients after hospitalization for heart failure. *Clin Pharmacol Ther*. 2010;88:548–554.
32. Harrell F, Lee KL, Mark DB. Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15:361–387.

33. Schwarz G. Estimating the dimension of a model. *Ann Stat*. 1978;6: 461–464.
34. Jones BL, Nagin DS. Advances in group-based trajectory modeling and an SAS procedure for estimating them. *Sociol Methods Res*. 2007; 35:542–571.
35. Liu X, Engel CC. Predicting longitudinal trajectories of health probabilities with random-effects multinomial logit regression. *Stat Med*. 2012;31:4087–4101.
36. Muthén B. Latent variable analysis: growth mixture modeling and related techniques for longitudinal data. In: Kaplan D, ed. *The Sage Handbook of Quantitative Methodology for the Social Sciences*. Thousand Oaks, CA: Sage Publications; 2004:345–368.
37. Hartigan JA, Wong MA. A k-means clustering algorithm. *Appl Stat*. 1979;28:100–108.
38. McCutcheon AL. *Latent Class Analysis*. Newbury Park, CA: Sage Publications; 1987.