Analysis of the North Carolina Long-Term Care Polypharmacy Initiative: A Multiple-Cohort Approach Using Propensity-Score Matching for Both Evaluation and Targeting

Troy K. Trygstad, PharmD, MBA, PhD[1,2]; Dale B. Christensen, RPh, PhD[3,4]; Steve E. Wegner, MD, JD[2,5]; Rob Sullivan, MD, MPH[1,6]; and Jennifer M. Garmise, PharmD[5]

[1]Community Care of North Carolina, Raleigh, North Carolina; [2]Department of Pediatrics, The University of North Carolina, Chapel Hill, North Carolina; [3]School of Pharmacy, The University of North Carolina, Chapel Hill, North Carolina; [4]School of Pharmacy, University of Washington, Seattle, Washington; [5]AccessCare Inc., A Community Care of North Carolina Network, Morrisville, North Carolina; and [6]Department of Community and Family Medicine, Center for the Study of Aging and Human Development, Duke University Medical Center, Durham, North Carolina

ABSTRACT

Background: The high cost and undesirable consequences of polypharmacy are well-recognized problems among elderly long-term care (LTC) residents. Despite the implementation of the 1987 Omnibus Budget Reconciliation Act, which requires pharmacist review of drug regimens in this setting, medical and drug costs for LTC residents have continued to increase.

Objective: This study evaluates the North Carolina Long-Term Care Polypharmacy Initiative, a large-scale medication therapy management program (MTMP) that combined drug utilization review activities with drug regimen review techniques.

Methods: This was a prospective records-based study that used a difference-in-difference model with both historical and nonintervention group controls. To ensure equivalence among subjects, propensity scoring was used to match study subjects from participating LTC facilities with comparison subjects from nonparticipating facilities. Residents with interventions were grouped for analysis by intervention type—retrospective only, prospective only, or dual type (residents with both prospective and retrospective interventions)—and by intervention stage—review, recommendation, and drug change—plus an all-inclusive "all types" grouping that aggregated groups by intervention type, for a total of 10 total cohorts.

Results: In the overall population of 5255 study subjects identified, a US $21.63 per member per month drug-cost savings was observed. Although only 1 of 10 cohorts had a change in the number of drug fills, substantial reductions in 2 of 5 types of drug alerts were observed in all 10 cohorts. A reduction in the relative risk for hospitalization (0.84 [95% CI, 0.71–1.00]) was observed in the cohort of residents receiving a retrospective review.

Conclusions: This Initiative suggests that an MTMP can be quickly launched in a large number of LTC facility residents to produce monetary drug-cost savings and improved health outcomes. Additionally, the evaluation of this program illustrates the utility of using propensity scoring techniques to target future intervention groups in a cost-effective manner. (*Clin Ther.* 2009;31:2018–2037) © 2009 Excerpta Medica Inc.

Key words: polypharmacy, drug utilization review, Medicaid, pharmacist, propensity scoring.

INTRODUCTION

The high cost and undesirable consequences of polypharmacy are well-recognized problems among elderly long-term care (LTC) residents.[1–4] Inappropriate polypharmacy can result in unnecessary monetary costs and undesirable patient outcomes such as adverse drug reactions, drug–drug interactions, and unnecessary hospitalization.[5–11]

To address these problems and decrease costs across the United States, the 1987 Omnibus Budget Reconciliation Act (OBRA) required that LTC phar-

Accepted for publication June 30, 2009.
doi:10.1016/j.clinthera.2009.09.006
0149-2918/$ - see front matter

macists and prescribers cooperate in prospective and retrospective drug regimen reviews (DRR) for elderly Medicaid patients. Despite these efforts, both medical care and drug costs have continued to increase.[12–14] The medication therapy management program (MTMP) provisions of the Medicare Modernization Act of 2003 (MMA) have a similar objective, and their impact have yet to be fully evaluated.

This paper describes a unique DRR intervention designed to minimize inappropriate polypharmacy in LTC facilities in North Carolina. It builds on 2 previously published studies that described The North Carolina Long-Term Care Polypharmacy Initiative (hereafter, "Initiative") interventions and resultant cost avoidance in detail.[15,16] These previous evaluations focused solely on outcomes of retrospective interventions and excluded patients having prospective interventions due to the suspected presence of selection bias.

The aim of the present study was to use a more rigorous analysis design to fully evaluate the downstream effects of pharmacist interventions on patient drug-related outcomes. We employed strict inclusion and exclusion criteria and created a propensity-matched comparison group of patients who did not receive the intervention. Unlike previous evaluations, we included all patients receiving interventions, regardless of the type. More specifically, the study aims were to: assess postintervention changes in drug utilization, drug cost, and prevalence of drug alerts and hospital admissions; compare and contrast the effects of prospective and retrospective interventions, both separately and in combination, on drug-related processes and outcomes; conduct a substrata analysis by propensity score to determine differential downstream results based on the likelihood of selection for intervention; and describe policy implications of providing MTMP services to elderly nursing facility patients.

METHODS

The Initiative was a cooperative effort between LTC consulting pharmacists and LTC providers to identify, test, and implement a method of value-added DRR that met the needs of the patient, pharmacist, and physician. The program was conducted under the auspices of Community Care of North Carolina (CCNC), a not-for-profit organization that provides primary care case management to North Carolina Medicaid recipients. CCNC operates through collaborative agreements with local community organizations and physician-group practices. The program was conducted prior to the passage of the MMA, and serves as a large-scale example of how a planned program of focused DRR for LTC patients, structured along the lines of the MTMP provisions of the MMA, can be comprehensively evaluated.

Consultant pharmacists performed targeted, value-added DRRs for selected Medicaid-dependent residents of LTC facilities in North Carolina during the routine monthly DRRs required by OBRA nursing facility guidelines. Consultant pharmacists were paid $12.50 for each review in addition to their own previously negotiated DRR rates already established with each facility. Drug claims data were used to create drug profiles that contained cost and quality-focused alerts for patients with ≥18 drug fills in the 90 days immediately preceding the intervention. Computer algorithms were used to screen profiles for 5 types of drug alerts. The first was related to potentially inappropriate medications in the elderly (Beers drug list),[17,18] the second was a suggestion for a therapeutic change to a more cost-effective drug (Prescription Advantage List [PAL], created by CCNC), the third was related to quality (Clinical Initiatives, created by an LTC pharmacist expert panel), the fourth was a duration alert based on specific drugs indicated only for acute or short-term use, and the fifth was a therapeutic duplication alert based on mechanism of action. These alerts were generated retrospectively from claims data and provided to the LTC consultant pharmacists for their retrospective reviews, together with each residents' most recent drug-claims profile. These profiles were comprehensive in nature and considered all drugs on a resident's profile, regardless of the presence of an alert.

The program also had a prospective component. As new medication orders came into the dispensing facility, a pharmacist could intervene and request a drug change using the same alert-targeting criteria developed for the retrospective, computer-generated drug profiles. Pharmacists were compensated US $6.50 for each prospective drug therapy problem identified and documented. Unlike retrospective reviews, compensation was given for each drug recommendation made. Some residents received only retrospective reviews and interventions, some received only prospective interventions, and some received both types.

This was a prospective records-based study that used both historical and nonintervention controls.

Comparison-group residents were drawn from nonparticipating LTC facilities. Propensity scoring was used to ensure the equivalence of study and comparison-group residents.

A total of 12 cohorts (3 × 4) were possible using this intervention and process matrix. We grouped residents based on intervention type (retrospective only, dual type, and prospective only) and 3 stages of treatment (review, recommendation, and drug change) plus an all-inclusive "all types" grouping that aggregated groups by intervention type. Of the 12 possible cohorts in the matrix, 2 involving prospective interventions at the review level were ignored since all patients with a prospective review received a recommendation by default, rendering those cohorts moot at the review level of treatment (**Figure 1**). Creation of the cohort matrix was necessary to parse out effect by intervention type as well as to establish a causal link with successive levels of treatment (review, recommendation, or drug change). Evaluation at each of these intervention levels was desirable because each stage represents a different intent-to-treat perspective that can guide future service-payment models.

Prescription drug records of all North Carolina nursing facilities were retrieved from Medicaid claims databases for the period of August 2002 through April 2003. This period encompassed the 90-day baseline, the 90-day intervention, and the 90-day postintervention periods to allow for a difference-in-difference (DID) with a comparison-group study method. Residents were required to be continuously eligible throughout the 90-day preintervention period, as well as the 90-day postintervention period. The 90-day postintervention period was chosen to balance a trade-off between a need to minimize potential dropouts in this elderly population and the need for a sufficient time horizon to measure effect. We excluded residents who were not continuously eligible for Medicaid during the study, had a hospitalization or an emergency department visit in the preintervention period, did not have a prescription filled in the first 35 days of the preintervention period or the last 35 days of the postperiod, or used a third-party payer other than North Carolina Medicaid.

Multifaceted and potentially substantial selection bias was anticipated. Potential sources of bias differed by the stage of treatment and the type of intervention, and were likely to include (among other sources): solicitation and subsequent self-selection of project sites at the facility level, selection of prospective action at the resident level, identification of actionable drug problems at the pharmacist level, and selective recommendation acceptance at the prescriber level.

Not only was selection bias presumed present, it was likely to be progressive throughout. Successive levels of intervention (review, recommendation, and drug change) might impose successive layers of bias (ie, facility self-selection, pharmacist selection, and then prescriber selection). These layers of bias were likely created sequentially as drug problems filtered through the initial screen to a pharmacist for review and then to a physician for acceptance or rejection. Because varying sources and degrees of presumed selection bias were anticipated, we formed propensity-matched

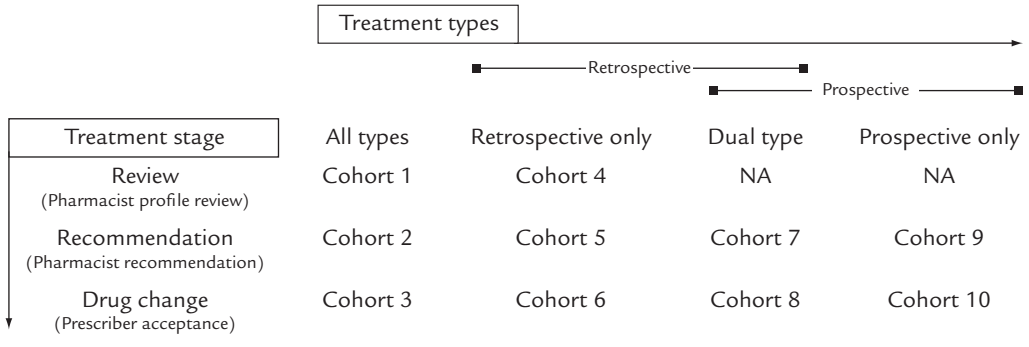

Figure 1. Cohort matrix. NA = not applicable because all residents with a prospective review received a recommendation by default.

cohorts of comparison-group residents in nonparticipating nursing facilities for each study cohort (**Figure 2**).

As of publication, no well-established criteria for developing propensity score models were known[19]; however, there is consensus that all potentially relevant covariates with higher-order terms and interacted terms be included.[20–23] Propensity scores are most useful when the relationship between baseline risk factors and treatment selection is not fully understood.[24] For these reasons, we included all variables that may have influenced treatment selection or response to the intervention. For example, the drug alerts likely prompted pharmacists to action. Any differences in the type or number of alerts between the intervention group and controls would have created a biased result. Thus, including all 5 drug-alert types in the propensity-scoring model improves similarity of the groups. Although variables such as age may have had a lesser influence, current practice is to include as many variables that may contribute to selection bias as possible.

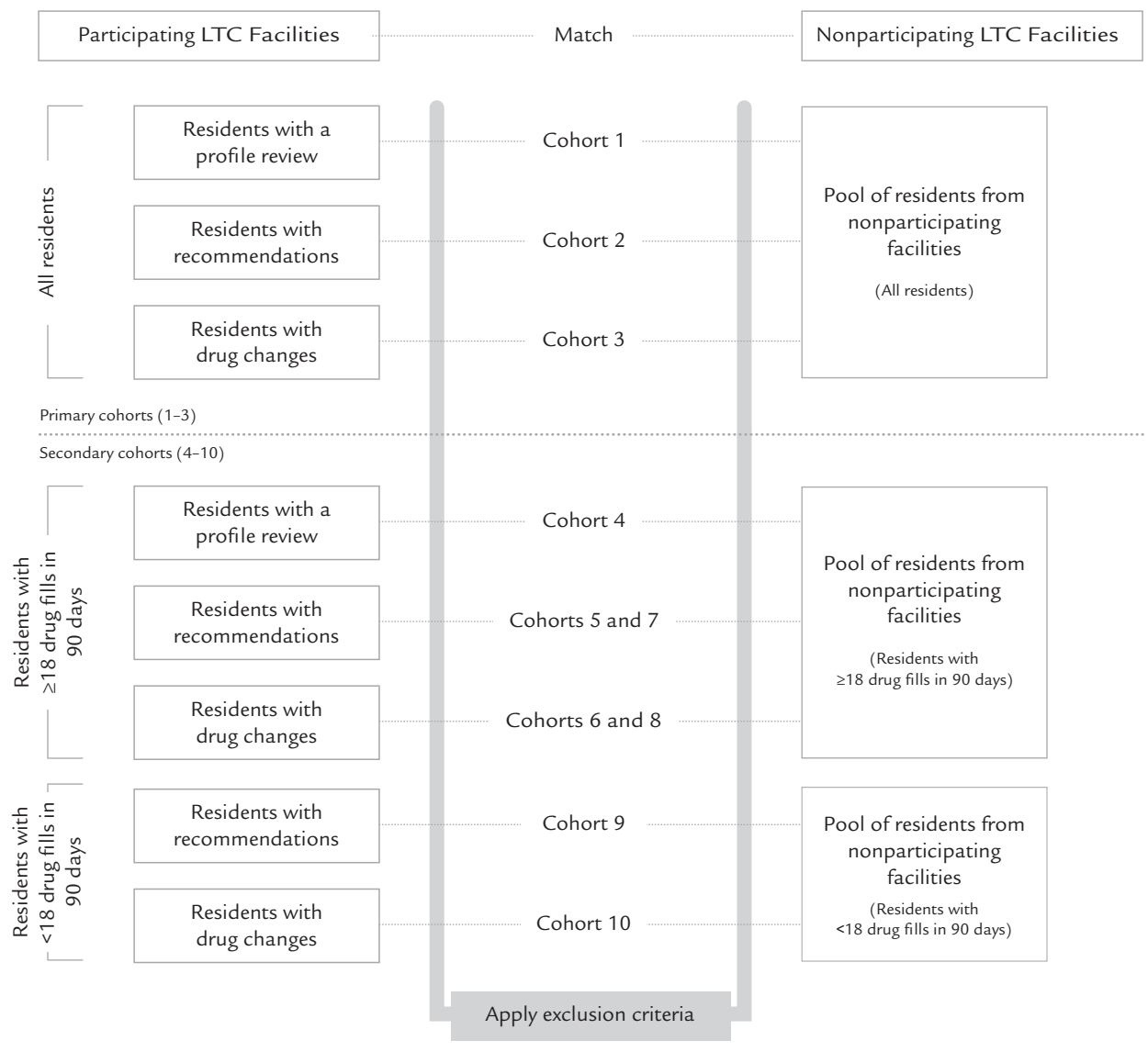

Figure 2. Propensity score–matching strategy. LTC = long-term care.

To avoid overmatching,[25] outcomes variables, such as posttreatment drug costs, were not included because they are dependent data elements used to observe for intervention effects. Requiring equivalence would result in a loss of attribution of intervention effect, which is the purpose of the analysis. The fully specified, fully interacted model is as follows:

Treatment selection = Age + $(\text{Age})^2$ + Race + Sex + Total no. of drug fills + Total preperiod drug cost + Total no. of alerts + No. of duplications + No. of Beers list drugs + Length of drug therapy alerts + No. of PAL alerts + No. of Clinical Initiatives alerts + $(\text{No. of supplications})^2$ + $(\text{No. of Beers list drugs})^2$ + $(\text{No. of drug therapy alerts})^2$ + $(\text{No. of PAL alerts})^2$ + $(\text{No. of Clinical Initiatives alerts})^2$ + (No. of drugs × No. of alerts) + (No. of drugs × Cost of drugs) + (No. of alerts × Cost of drugs) + Error term.

After propensity scoring was complete, we matched both study and potential comparison subjects using Mahalanobis metric matching to achieve balance among baseline characteristics. This method was chosen for its robust ability to determine the closest match and its amenability to matching with replacement.[26] A replacement method was necessitated by the existence of cohorts in which study subjects outnumbered potential comparison subjects. Nearest match without replacement methods work as well as replacement methods only when a sufficient number of "relevant comparison units" are available.[27] After matching, we tested for balance using the *t* test for continuous data and the χ^2 test for categorical data.

To visualize the distribution of bias in the 10 cohorts, a density graph was created and centered on the mean probability of treatment. This representation of bias can be inspected against a pool of comparison subjects to determine whether matching is needed a priori and to check for equivalence across heterogeneous groups postmatch. Bias is represented by diverging curves. Where the lines overlap, no bias exists with respect to the observed variables in the propensity-scoring model.

After the successful match, downstream outcomes were assessed. To assess changes in drug use, we calculated DID amounts for total drug cost, number of prescription fills, and number of potential drug therapy problem alerts (PAL, Clinical Initiatives, Beers list, therapeutic duplication, and length of drug therapy alerts). To assess changes in downstream health outcomes, we created a dichotomous variable that identified residents having a hospitalization in the postperiod.

We used Wilcoxon signed rank testing to assess between-group changes in total drug cost, number of prescription fills, and alert rates. We used the χ^2 distribution and the Fisher exact test for between-groups testing of relative risk (RR) estimates for hospitalization. All statistical analyses were conducted with Stata statistical software version 9 (StataCorp LP, College Station, Texas) and $P \leq 0.05$ was considered statistically significant (2-tailed). Mahalanobis matching was performed using PSMATCH2.[28] Finally, we stratified residents based on their likelihood of receiving an intervention (their propensity score) and then calculated groupwise before–after reductions in total drug cost. This was done for the all-inclusive primary cohort 1 (all residents receiving the intervention regardless of type and level) to determine whether the cost–benefit ratio of the intervention differed by the likelihood of receiving an intervention. This subgroup analysis was performed by evaluating study and comparison subjects by propensity score quintiles, with the lower quintiles having residents with characteristics less predictive of intervention and the upper quintiles having resident characteristics more predictive of intervention.

RESULTS

Of the 384 LTC facilities billing units in North Carolina, 253 chose to participate. Prescription drug records of all North Carolina nursing facilities were retrieved from Medicaid claims databases. **Figure 3** shows reductions in the sample size as exclusion criteria were applied. Of the 8087 residents who received an intervention, 7298 were continuously eligible throughout the study period, resulting in a 9.8% loss to eligibility. Of the 7298 continuously eligible, 5917 had a prescription fill in the last 35 days of the postperiod, representing a 17.1% loss to follow-up. When the exclusion criteria (hospitalization, emergency room visit, lack of prescription claims, or residence in a nursing facility) were applied, 5255 residents remained for the analysis, resulting in an additional 8.2% reduction in sample size. Of the 5255 patients remaining for analysis, 3618 received a recommendation and 2517 had a drug change resulting from that recommendation (**Figure 4**). The 35% loss to the exclusion criteria was nearly identical (0.04% difference) to the loss experienced in the comparison-group facilities.

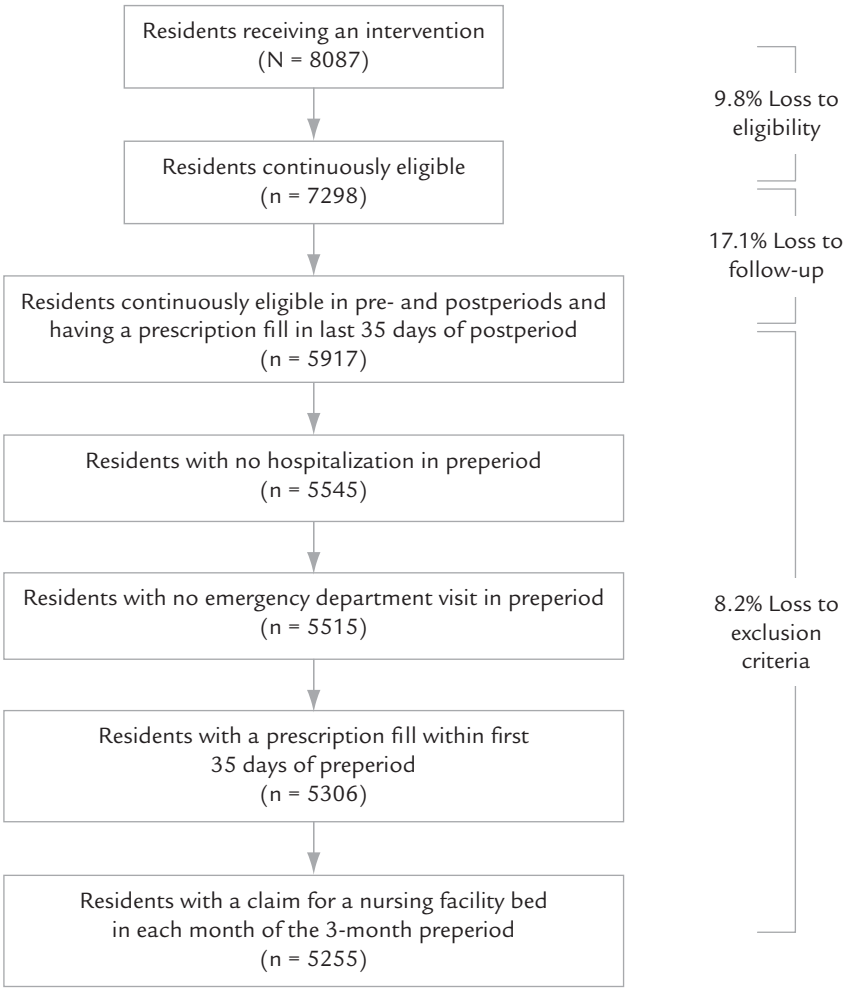

Figure 3. Resident disposition and application of exclusion criteria.

Pre- and postmatch absolute percentage bias is reported in **Table I**. Prospective interventions (including dual type) produced substantial prematch bias. As anticipated, successive stages of intervention (review, recommendation, and drug change) produced successively increasing prematch bias. The greatest bias existed when all intervention types were considered as a whole, including patients with <18 drug fills in the previous 90 days (the screening criteria that triggered a retrospective review). Postmatching bias was consistent for all cohorts and ranged from 5.06 to 7.97, which is a desirable range for absolute percentage bias.[29]

Figures 5 and **6** illustrate pre- and postmatching density distributions, respectively, with regard to the likelihood of receiving treatment. Visual inspection confirms that propensity-scored matching was warranted a priori for most of the cohorts and that postmatch balance was achieved. Baseline (prematching) and postmatching characteristics of cohort 1 as well as their pre- and postmatch percent biases are presented in **Table II**.

The overall impact on drug costs, utilization, and clinical alerts is reported in **Table III**. All results are based on intent-to-treat analyses for the various primary and secondary cohorts. Statistically significant drug-cost savings were observed in 9 of the 10 cohorts, including the primary cohort (1), which had a 4.4% cost reduction and a $21.36 per member per

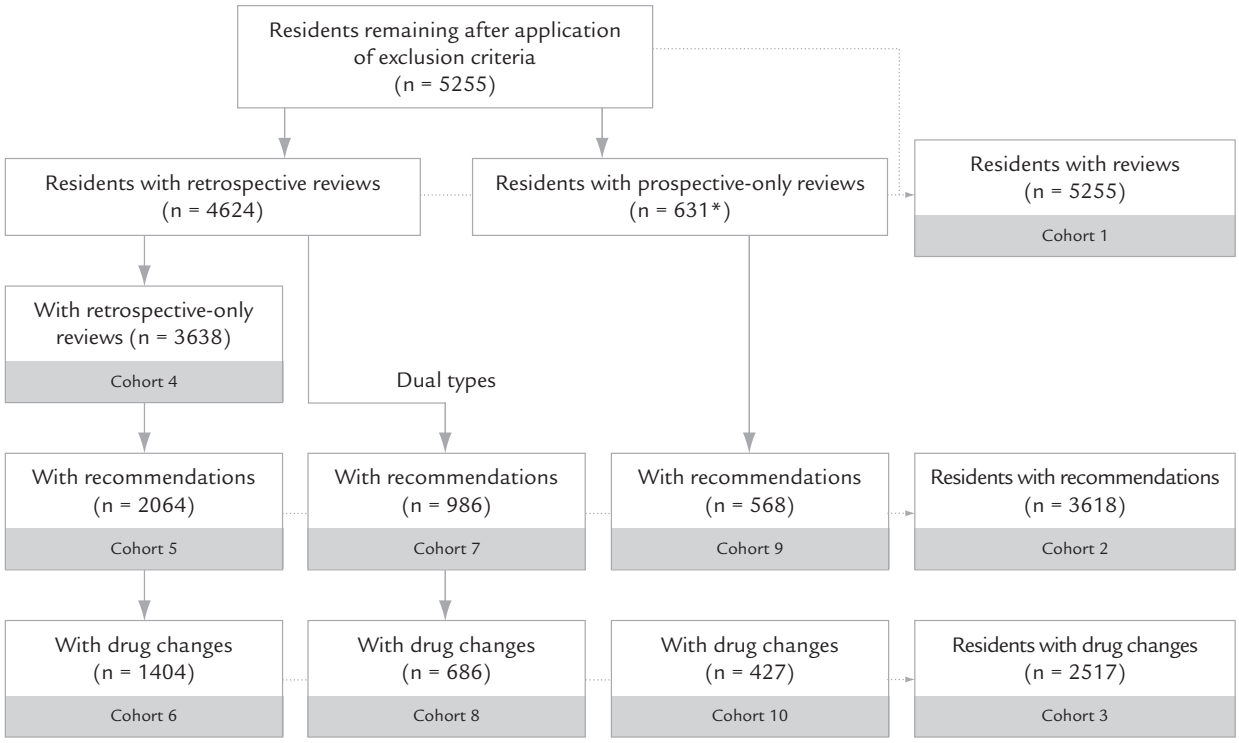

Figure 4. Resident disposition according to reviews, recommendations, and drug changes. *Sixty-three residents had a prospective review with no recommendation under a special circumstance; therefore, they were excluded from the subgroup analysis.

Table I. Mean absolute percentage bias reduction by intervention stage and type.

	All Types		Retrospective Only		Dual Type		Prospective Only	
Stage	Prematch	Postmatch	Prematch	Postmatch	Prematch	Postmatch	Prematch	Postmatch
Review	24.89	6.53	4.92	5.06	NA	NA	NA	NA
Review and recommendation	28.90	7.83	6.55	6.97	12.75	7.48	17.66	5.78
Review, recommendation, and drug change	30.18	7.97	8.61	6.80	16.36	7.64	18.93	6.67

NA = not applicable.

month (PMPM) savings ($P < 0.05$). Among the other cohorts of the same intervention type there was a 6.3% reduction ($P < 0.001$) at the recommendation level and a 7.8% reduction ($P < 0.001$) at the drug-change level. The savings increased at intervention types and levels successively closer to the drug-change stage (ie, moving from [review] to [review and recommendation] to [review, recommendation, and drug change]). The largest saving was at the drug-change level for residents with retrospective-only interven-

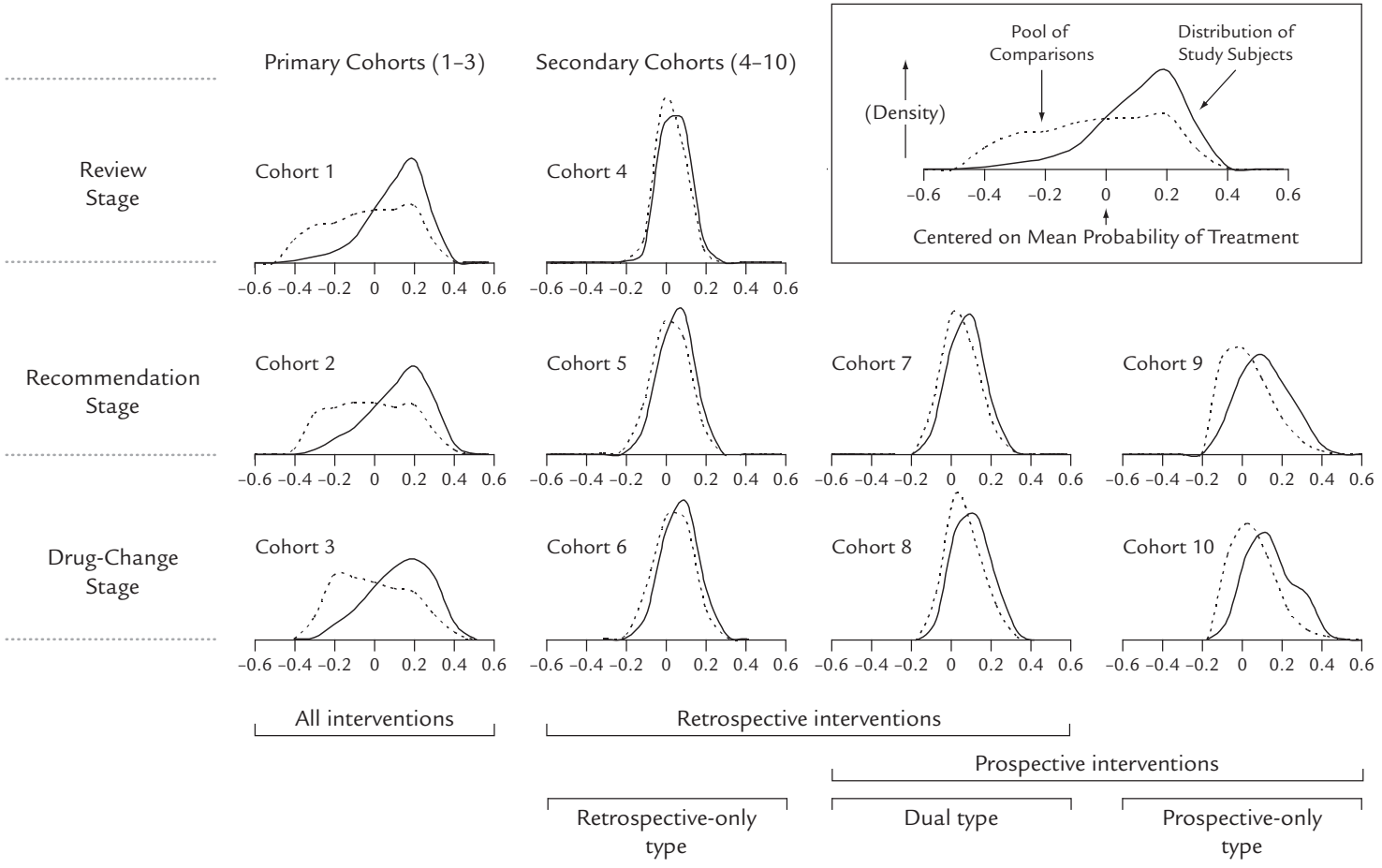

Figure 5. Distribution of propensity scores between study and comparison subjects (prematch).

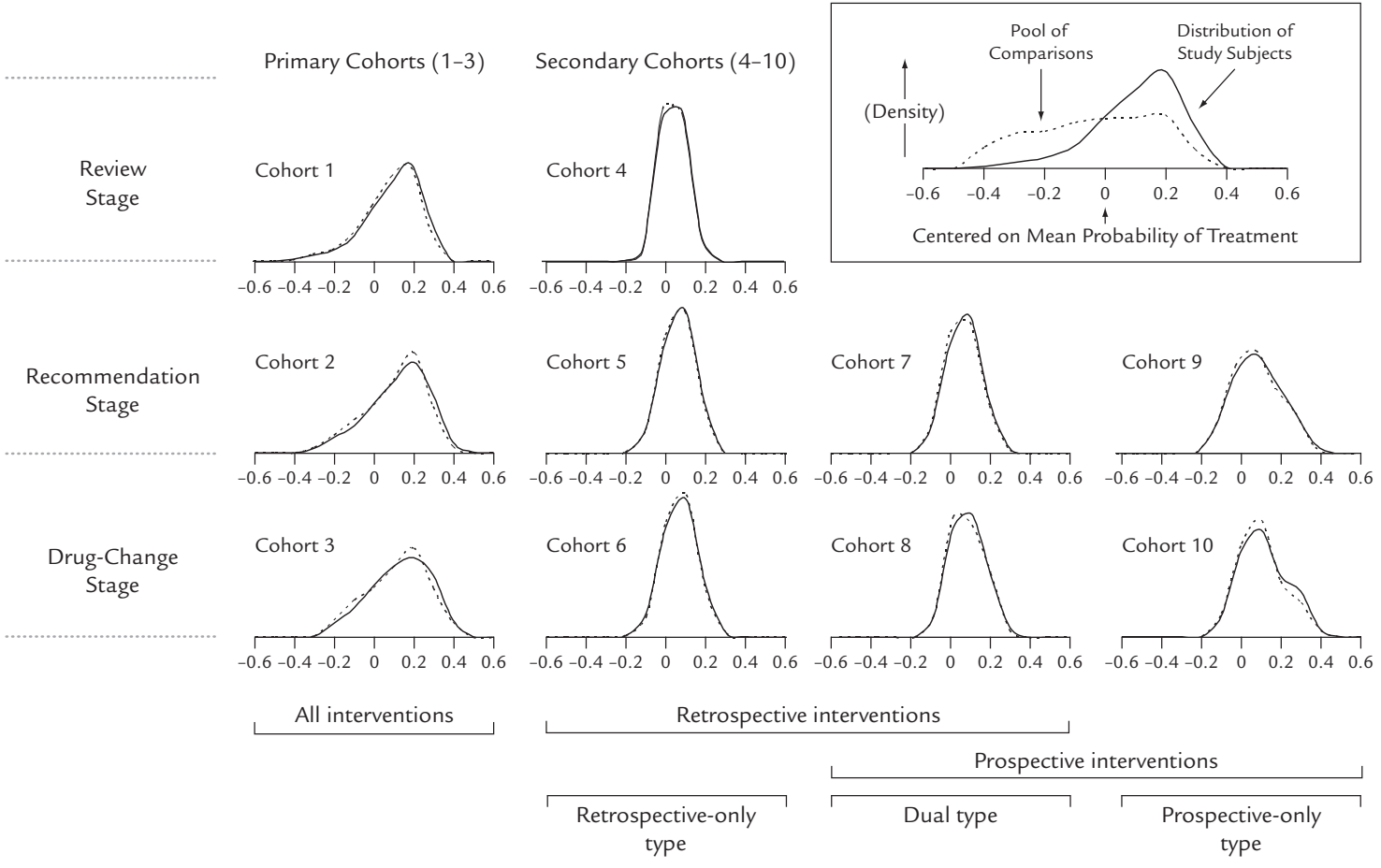

Figure 6. Distribution of propensity scores between study and comparison subjects (postmatch).

Table II. Pre- and postpropensity score–matched baseline characteristics and bias with reductions in cohort 1 (all intervention types, review stage).*

Variable/Sample	Mean Treated	Mean Comparison	% Bias	% Reduction in Bias	*t*	*P*
Age, y						
Prematch	77.6	79.4	–14.1	–	–6.62	<0.001
Postmatch	77.6	78.5	–7.1	–50	–3.83	<0.001
(Age)^{2}†						
Prematch	6175	6455	–15.7	–	–7.38	<0.001
Postmatch	6175	6286	–6.2	–61	–3.35	0.001
Nonwhite race						
Prematch	32.2%	24.9%	16.0	–	7.48	<0.001
Postmatch	32.2%	30.5%	3.6	–78	1.79	0.074
Female sex						
Prematch	75.1%	78.8%	–8.7	–	–4.05	<0.001
Postmatch	75.1%	76.2%	–2.6	–70	–1.30	0.195
Total no. of drugs						
Prematch	26.9	20.4	57.5	–	27.22	<0.001
Postmatch	26.9	25.6	11.0	–81	6.11	<0.001
Total amount paid						
Prematch	1442	1088	27.1	–	12.19	<0.001
Postmatch	1442	1340	7.8	–71	4.04	<0.001
Total no. of alerts						
Prematch	9.37	6.73	48.1	–	22.57	<0.001
Postmatch	9.37	8.86	9.4	–81	4.98	<0.001
No. of duplication alerts						
Prematch	4.47	3.04	41.9	–	19.55	<0.001
Postmatch	4.47	4.14	9.8	–77	5.10	<0.001
(No. of duplication alerts)^{2}†						
Prematch	32.4	20.1	25.4	–	11.82	<0.001
Postmatch	32.4	27.3	10.6	–58	5.57	<0.001
No. of Beers list alerts						
Prematch	0.686	0.522	18.4	–	8.55	<0.001
Postmatch	0.686	0.645	4.6	–75	2.29	0.022
(No. of Beers list alerts)^{2}†						
Prematch	1.35	0.98	12.5	–	5.78	<0.001
Postmatch	1.35	1.20	5.1	–59	2.53	0.012
No. of PAL list alerts						
Prematch	1.47	1.14	27.9	–	13.03	<0.001
Postmatch	1.47	1.42	3.5	–87	1.82	0.069
(No. of PAL list alerts)^{2}†						
Prematch	3.63	2.60	20.6	–	9.53	<0.001
Postmatch	3.63	3.31	6.4	–69	3.19	0.001

(continued)

Table II (continued).

Variable/Sample	Mean Treated	Mean Comparison	% Bias	% Reduction in Bias	*t*	*P*
No. of Clinical Initiatives alerts						
Prematch	2.60	1.89	42.1	–	19.72	<0.001
Postmatch	2.60	2.51	5.5	–87	2.87	0.004
(No. of Clinical Initiatives alerts)^2						
Prematch	9.62	6.33	29.7	–	13.75	<0.001
Postmatch	9.62	8.76	7.8	–74	3.87	<0.001
No. of consider length alerts						
Prematch	0.15	0.135	3.3	–	1.55	0.121
Postmatch	0.15	0.143	1.6	–51	0.80	0.423
(No. of consider length alerts)^2						
Prematch	0.23	0.187	4.7	–	2.17	0.030
Postmatch	0.23	0.218	1.4	–71	0.66	0.511
Total no. of alerts × total no. of drugs						
Prematch	293.7	187.8	37.7	–	17.54	<0.001
Postmatch	293.7	264.3	10.5	–72	5.45	<0.001
Total amount paid × total no. of drugs						
Prematch	44,936	29,281	22.5	–	10.05	<0.001
Postmatch	44,936	39,467	7.9	–65	4.04	<0.001
Total amount paid × total no. of alerts						
Prematch	16,183	10,170	23.9	–	10.69	<0.001
Postmatch	16,183	14,049	8.5	–65	4.33	<0.001

PAL = Prescription Advantage List.

*Sample sizes: study group, 5255; unmatched comparison group, 3801; matched comparison group, 5255. Absolute % bias: prematch, 24.89; postmatch, 6.53; change in bias: –74%. Pseudo-R^2: prematch, 0.105; postmatch, 0.008.

†Interacted variables are imputed into the model by squaring (multiplying the value by itself) and serve to represent nonlinear relationships between the variable and the dependent variable.

tions, with a reduction of 8.0% ($41.96 PMPM; $P < 0.05$). For residents having recommendations, the mean number of baseline drug fills was numerically higher (29.3 drug fills) than the prospective reviews (12.5 drug fills), which were initiated based on a single drug with an alert. Subsequently, the mean number of alerts (1.33 vs 4.52) and, ultimately, opportunities for alerts were greater in the retrospective (targeted, retrospective) group versus the prospective group (event driven).

The number of prescriptions remained unchanged in 9 of the 10 cohorts, with an increase of 2.6% ($P = 0.05$) at the recommendation level for residents who had both a retrospective review and a prospective review. There were no other significant changes in the number of prescription fills among all other cohorts and interventions.

There were significant reductions in the incidence of PAL alerts for all 10 groups. In the primary cohort

Table III. Outcomes of Initiative activities by cohort.

Alert	All Intervention Types Combined			Retrospective Only			Dual Type		Prospective Only	
	Rev (Cohort 1; n = 2178)	Rev and Rec (Cohort 2; n = 1845)	Rev, Rec, and DC (Cohort 3; n = 1496)	Rev (Cohort 4; n = 1455)	Rev and Rec (Cohort 5; n = 1121)	Rev, Rec, and DC (Cohort 6; n = 906)	Rev and Rec (Cohort 7; n = 690)	Rev, Rec, and DC (Cohort 8; n = 527)	Rev and Rec (Cohort 9; n = 429)	Rev, Rec, and DC (Cohort 10; n = 344)
Drug cost, US$										
Δ Mean DID	−64.09*	−91.94*	−114.15*	−62.39*	−91.58*	−125.89*	−37.18	−83.84*	−110.83*	−120.15*
%Δ Mean DID	−4.4	−6.3	−7.8	−4.1	−5.9	−8.0	−2.3	−5.0	−15.0	−16.4
Δ PMPM DID	−21.36*	−30.64*	−38.05*	−20.80*	−30.52*	−41.96*	−12.39*	−27.95*	−36.94*	−40.05*
No. of drug fills										
Δ Mean DID	0.05	0.19	−0.09	−0.04	0.05	−0.46	0.81*	0.41	0.16	0.15
%Δ Mean DID	0.17	0.7	−0.34	−0.12	0.16	−1.6	2.6	1.3	1.3	1.2
PAL list alerts										
Δ Mean DID	−0.28*	−0.35*	−0.44*	−0.27*	−0.36*	−0.49*	−0.36*	−0.47*	−0.30*	−0.38*
%Δ Mean DID	−19.2	−21.7	−26.6	−18.1	−21.3	−27.9	−20.3	−25.6	−30.6	−37.7
Clinical Intiatives alerts										
Δ Mean DID	−0.25*	−0.25*	−0.31*	−0.30*	−0.30*	−0.37*	−0.22*	−0.31*	−0.17*	−0.24*
%Δ Mean DID	−9.6	−8.9	−10.9	−11.6	−10.4	−12.3	−7.0	−9.6	−10.3	−13.9
Beers list										
Δ Mean DID	−0.012	−0.018	−0.03	−0.03	−0.21	−0.033	−0.028	−0.41	−0.042	−0.037
%Δ Mean DID	−1.7	−2.6	−4.2	−4.2	−2.8	−4.5	−3.3	−4.4	−16.4	−14.3
Length of therapy alerts										
Δ Mean DID	−0.01	−0.007	−0.01	−0.008	0	−0.009	−0.016	−0.022	−0.03	−0.023
%Δ Mean DID	−6.4	−5.2	−7.9	−5.2	0	−6.6	−10.5	−15.5	−36.2	−27.8
Therapeutic duplication alerts										
Δ Mean DID	0.17*	−0.067	−0.104	−0.333	−0.167	−0.193	0.116	0.016	−0.035	−0.087
%Δ Mean DID	−3.8	−1.5	−2.4	−6.9	−3.5	−3.9	−2.3	−0.3	−2.3	−5.6

Rev = review; Rec = recommendation; DC = drug change; n = number of matched pairs; DID = difference-in-difference; PMPM = per member per month; PAL = Prescription Advantage List.
*$P < 0.05$ (2-tailed t test).

(1), these alerts had a DID reduction of 19.2% ($P < 0.001$) at the review stage for all intervention types. The reductions were most prominent for residents receiving prospective-only–type interventions, where a 37.7% reduction ($P < 0.001$) was found for residents with a drug change. As with drug cost reductions, the percentage of PAL alert reductions increased at successive intervention levels. The number of PAL alerts decreased in both the study and comparison groups. For cohort 1, a 28.0% reduction was observed versus a 9.8% reduction in the comparison group, resulting in a DID reduction of 19.2% ($P < 0.001$).

There were also significant DID reductions in the number of Clinical Initiatives alerts in all cohorts, including the primary cohort (1) (–9.6%; $P < 0.001$). In absolute terms, the reduction per resident in Clinical Initiatives alerts was 0.25 (–9.6%; $P < 0.001$), 0.25 (–8.9%; $P < 0.001$), and 0.31 (–10.9%; $P < 0.001$) for the review, recommendation, and drug-change levels, respectively, for all intervention types. Approximately one quarter of all residents who had a review and one third of those having any drug change had the net effect of 1 Clinical Initiatives list drug changed in their regimen.

There were no statistically significant reductions in Beers list alerts or length of drug therapy alerts.

A reduction in the RR of hospitalization was observed in residents receiving retrospective-only–type reviews (RR = 0.84; $P = 0.04$; 95% CI, 0.71–1.00), although the primary group had a point estimate of RR = 0.87 ($P = 0.066$; 95% CI, 0.75–1.01) and was statistically significant at a $P = 0.066$ level. Point estimates for the remaining 8 cohorts had an RR result <1.0, but were not statistically significant at a $P < 0.05$ level (**Figure 7**).

Figure 8 presents cost savings by quintile. No statistically significant drug-cost savings were found in the first 2 quintiles. Ninety-day drug-cost savings were significant in quintiles 3, 4, and 5 (–$83, –$91, and –$149). The mean DID cost reductions increased proportionally as the probability of receiving treatment increased.

DISCUSSION

The Initiative was a success. Although we only assessed the first phase of the Initiative in this paper, the 3-year project (2002–2005) included a pilot phase and 3 implementation phases with 19,144 LTC residents receiving intervention. Over the course of the project, pharmacists made 17,545 recommendations that led to >10,000 drug changes and an estimated $9 to $12 million in drug-cost savings annualized.[30]

In this paper, we focused on phase I and evaluated cost and quality outcomes for all study participants, regardless of intervention type or level of treatment. The results herein corroborate previous findings that included only a subset of patients receiving retrospective reviews.[15,16] This more sophisticated approach used propensity scoring, allowing for the assessment of residents receiving interventions that were excluded from previous analyses due to a high likelihood of selection bias.

The equivalence of the review and recommendation treatment levels for retrospective-only interventions was not enhanced by propensity-score matching, validating 2 previous studies[15,16] that focused solely on this intervention type, but without using a matching technique. For all other groups, propensity scoring improved equivalency, confirming the exclusion of these groups from previous analyses due to suspected selection bias.

The cohort matrix was used because payment models for this DRR service might reflect different levels of treatment and types of intervention. Some programs pay for services at the review level (which is the most prevalent model [similar to the current physician-payment model for evaluation and monitoring]), some pay at the recommendation level (in which a drug problem is presumed to exist and is acted on [similar to a physician payment for a procedure], but may not produce an effect), and some pay at a result level (requiring an actual change [an outcome] to occur for payment). Although the findings from the present study are a guide and give relative value to the different levels of intervention, program administrators should be cognizant of the variations in intervention focus, implementation, and follow-through as they relate to the types of recommendations made, the residents for whom they were made, and the level of physician acceptance. Perhaps more important than the results herein, we believe that we present a robust approach to evaluating a large-scale, diverse intervention involving multiple sites, pharmacists, and prescribers.

We found that targeted DRR interventions led to drug-cost savings of $21.36 PMPM, an amount that exceeded the 1-time payment of $12.50 pharmacists were paid for each review. Earlier studies that were

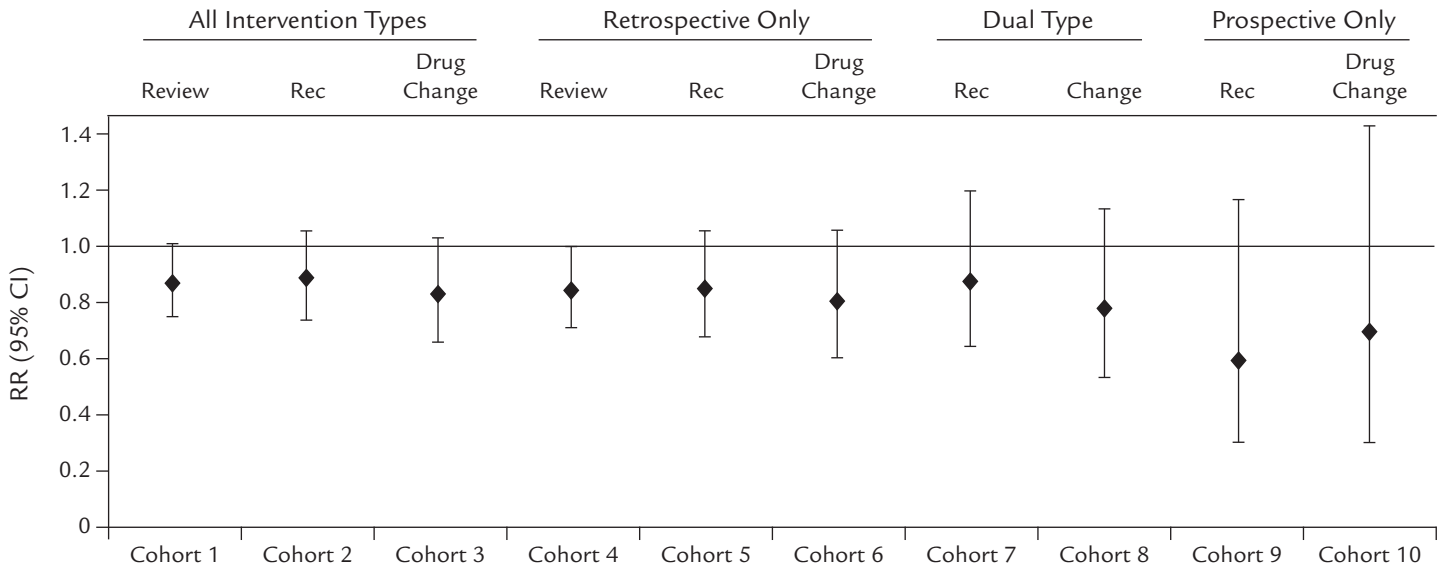

Figure 7. Relative risk (RR) for hospitalization resulting from intervention. Rec = recommendation. *Statistically significant ($P = 0.04$; Fisher exact test).

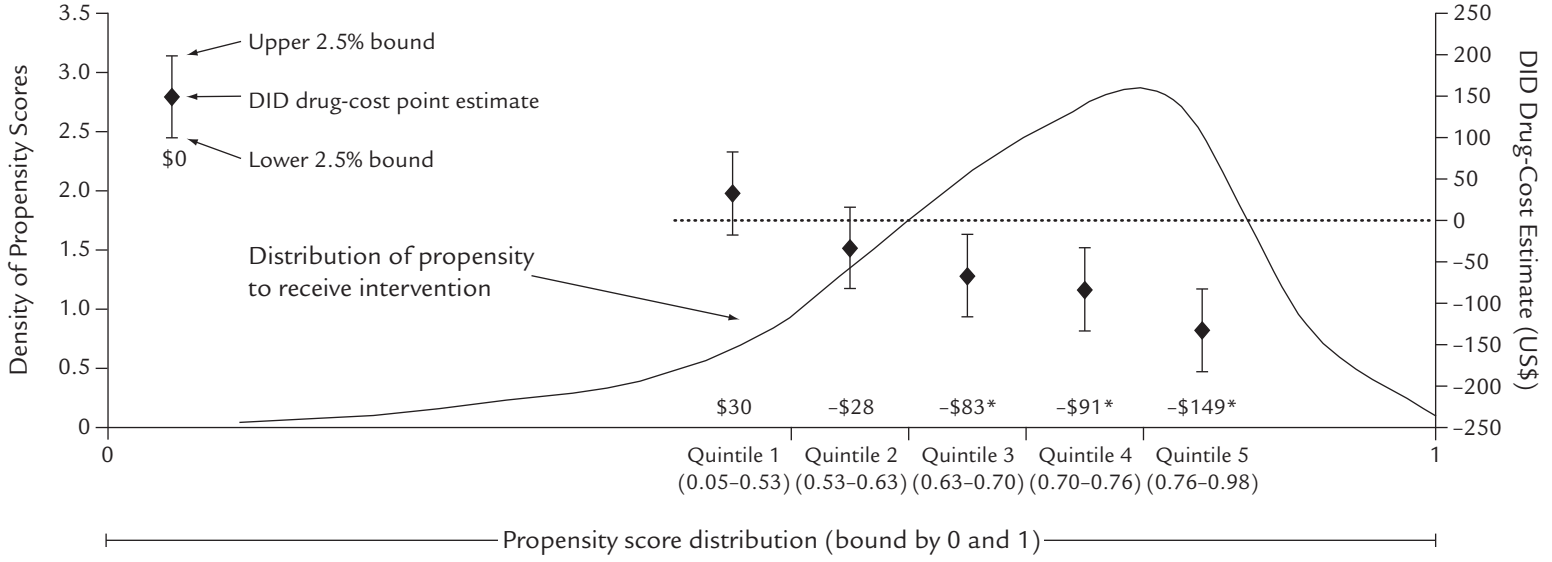

Figure 8. Drug-cost reduction by propensity-score quintile. DID = difference-in-difference. *Statistically significant (all, $P < 0.001$) drug-cost reduction (90-day postperiod).

limited to residents with retrospective-only–type interventions reported an average drug-cost reduction of $30.33 and $19.04 PMPM.[15,16] In other words, the intervention paid back its service-fee cost in the first month, although downstream savings were likely to accrue past the 90-day postperiod evaluation. The results from all 3 studies suggest that savings were sufficient to consider a higher level of payment to pharmacists, and/or paying prescribers for their participation.

Mean DID cost reductions increased proportionally as the probability of receiving treatment increased. This finding has important implications for MTMP administrators because it confirms that there are key patient-level characteristics that predict the success of pharmacist-led interventions. It suggests that an MTMP administrator or a payer such as Medicare may be able to determine the threshold of recipient characteristics that result in a net cost savings or quality improvement. For general application, a program administrator might propose that a pilot be conducted involving a sufficient number of subjects to provide for a logistic regression that is efficient (likely where $n > 1000$), then use the comparison group subjects' propensity scores along with the resulting cost or quality outcomes from the pilot group to target the larger population when the production mode of the program ensues. In this application, the findings from the pilot are applied to more efficiently target the larger population, since the program administrator now knows which pilot subgroups will have a positive return on investment. If we use these Initiative results as an example of this general application, any resident in a comparison LTC facility with a propensity score >0.63 in the cohort 1 model (**Figure 9**) would produce a positive return on investment based on these results and should be targeted for the next phase of the project. Note that the threshold propensity score (0.63) is model, cohort, and program specific.

This study used a robust case-matching technique to account for selection pressures in every aspect of the activity and reports on drug costs and health outcomes. Additionally, it was a statewide effort that affected the majority of recipients in the LTC setting, establishing a level of generalizability that is rare in intervention studies with pharmacists.

Nonetheless, this study was limited by the availability of only 3 months of postintervention data and 1 insensitive health outcome indicator, a hospitalization event. We did not address the potential for substitution effects, where nondrug costs may increase in response to drug-cost minimization efforts, though the hospital event data indicate a favorable global health outcome. Studies with longer time horizons and more sensitive and specific quality metrics are warranted. Yet, these studies remain elusive given the complexity of comorbid health conditions and the very high rate of patient attrition common to all LTC settings.

As in most practice-based, "real-world" interventions, it was not possible to draw a true random sample of patients across nursing facilities, physicians, or consultant pharmacists. In reality, LTC physicians and pharmacists often provide overlapping services to many of the same LTC facilities, making it impossible to avoid confounding. Additionally, groups of pharmacists are often clustered through consulting organizations that serve multiple nursing facilities and operate under a common ownership structure. Fortunately, participation tended to follow in accordance with pharmacists, physicians, and common ownership structures, thereby limiting overlap among study and comparison groups. Any overlap that did occur would serve only to reduce the reported effect, causing an underestimation of effect rather than an overestimation.

To address resident-level selection, we used propensity scoring because it offers a powerful method to reduce bias that, in some cases, performs better than randomization.[31] The purpose of this type of analysis is to match participants with comparable nonparticipants based on the likelihood of receiving an intervention (ie, selection) in an attempt to gain equivalence of groups. However, propensity scoring produces a valid result only if all relevant aspects of treatment selection and baseline risk are contained in the model. As with all nonrandomized, quasiexperimental methods, there is the possibility of the existence of a set of unobserved variables (not included in the propensity-scoring model) that are correlated with the outcome but were represented disproportionately in intervention and comparison groups. The model is limited to the baseline characteristics of the residents. Not contained in the model, but potentially present, were endogenous differences between participating and nonparticipating pharmacists, physicians, and/or nursing facilities. We believe, however, that the 20 covariate treatment selection model we used was robust and representative of most, if not all, selection pressures.

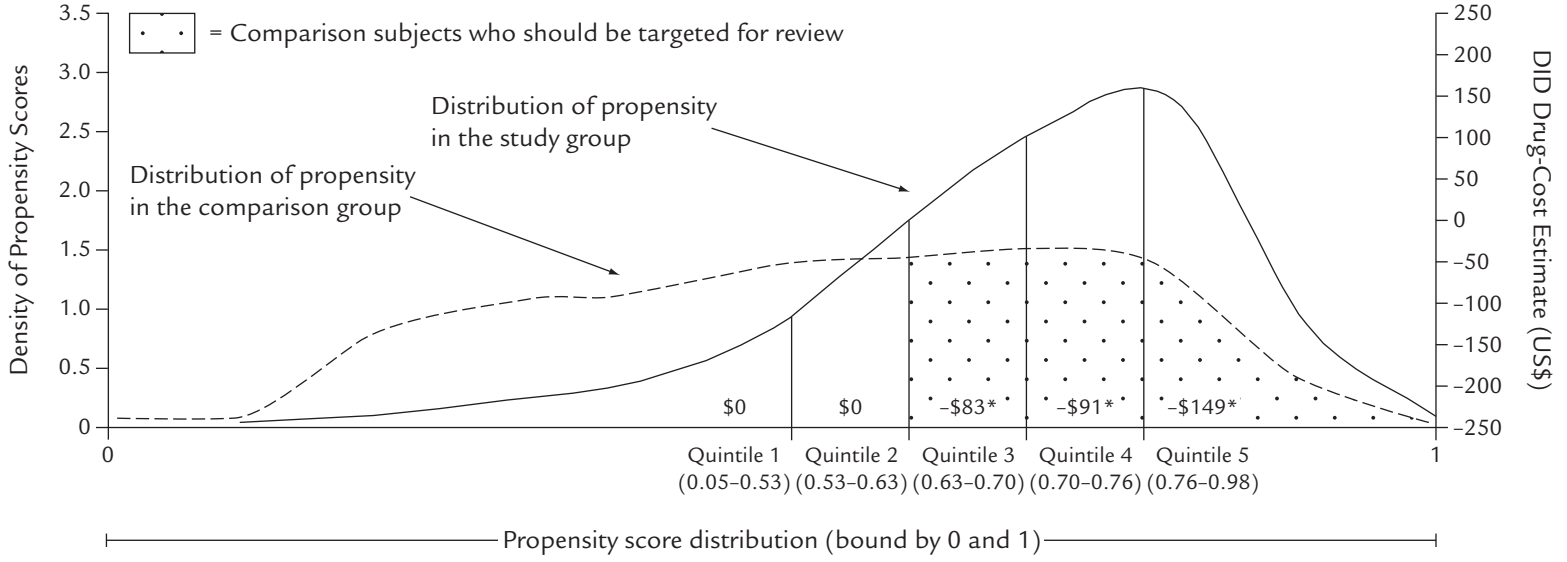

Figure 9. Example of targeting strategy for future intervention activities. DID = difference-in-difference. *Statistically significant (all, $P < 0.001$) drug-cost reduction (90-day postperiod).

Unlike more traditional approaches that match from a pool of comparison residents that are eligible for intervention, we matched against a pool of residents in nonparticipating homes. This has the effect of mitigating selection bias imposed by any unobserved variables that may exist since the comparison pool (nonparticipating facilities) did not have selection pressures imposed (at the resident level).

We chose a replacement method that matches study subjects with its nearest neighbor regardless of the frequency with which it has been replaced. Replacement of comparison subjects might be unsettling to the lay reader, but is necessary to match study participants with a sufficient number of comparable observations that are a close match. To date, no consensus or thorough review of matching techniques exist, although most researchers imply that choice of method depends on a variety of factors specific to each study.[27,32–34] Regardless, most replacement methods have been shown to be robust in practice,[35] especially when evaluating interventions,[35,36] and when the degree of overlap between treatment and comparison groups is large.[27]

Using administrative claims data to assess intervention studies has several inherent limitations, including the assumption that a drug claim represented a drug dispensed and consumed. In an LTC population, this appears to be a reasonable assumption. Although we eliminated patients who had a third-party insurance source, it is possible that patients might have acquired drugs from other sources. More to the point, this study takes a payer perspective, and paid claims are the most meaningful measurement. Further, participants in this study and comparison groups were equally subjected to the same limitations, further elucidating the advantages of a DID model of evaluation.

LTC resident attrition in North Carolina was unchanged for several years before the study (36% per year), hampering an evaluation of the persistence of intervention effects and damaging attempts for a rigid intent-to-treat approach. While we would have preferred a 6- to 12-month follow-up period, the ability to generalize these results would have suffered due to the possible existence of heterogeneous effects associated with persistence and survival. Therefore, we chose a 3-month follow-up period, reasoning that most of the impact of a drug therapy change should be seen within this time window. Extrapolation of findings beyond the 3-month follow-up must be done with caution, although health outcome measures such as rates of hospitalization would likely have been more sensitive with a longer time horizon.

The only other concurrent comparative program that has evaluated health outcomes is the Fleetwood Project.[37] This multiphase, multisite, multiyear study of drug therapy and tolerability issues in LTC settings emphasizes prospective interventions, face-to-face patient contact, and assessment of high-risk patients. In phase III of that project, 26 nursing facilities generated 2118 interventions for 4272 residents; 89% of dispensing pharmacists' recommendations and 55% of consultant pharmacists' recommendations were accepted.

The present study existed within the context of a preexisting retrospective DRR requirement for all residents of nursing facilities whose expenses were covered by federal health benefit programs, where comparison subjects received the standard of pharmaceutical care per OBRA regulations. We examined the marginal impact of a value-added DRR that differed from the usual DRR, in that it focused on patients who were determined by computer-based criteria to be at risk for an avoidable drug-related problem or cost-effective drug substitution. Although DRR programs are legislatively and inherently different from drug-utilization review (DUR) programs, which monitor drug use across Medicaid populations, and prior authorization (PA) programs, which require prescriber-affirmed criteria before coverage, this Initiative did result in high numbers of therapeutic interchanges (switching to drug products with presumed therapeutic equivalency). DUR or PA programs with prevalent therapeutic interchange have been criticized for producing unintended consequences,[38–44] largely in the domain of nondrug costs or outcomes. The results from this analysis suggest that there may be a reduction in hospitalizations, or a null effect, on nondrug costs and outcomes. This intervention did not include PAs, reimbursement limits, or other restrictions on drug selection and was completely voluntary in all aspects. It also focused on cost and quality of pharmaceutical use among residents of LTC facilities.

CONCLUSIONS

The North Carolina Long-Term Care Polypharmacy Initiative, a unique large-scale intervention and successful MTMP that meets MMA Part D requirements, demonstrated that combining DUR activities with DRR is valuable and feasible. It also demonstrated

that a computerized alert system that identifies potential drug therapy problems can be quickly launched in large numbers of LTC patients. It was used to enhance pharmacists' ability to review patient profiles and recommend changes that led to large monetary savings while maintaining effective drug therapy outcomes. Propensity scoring is both an effective evaluation tool and prognostic indicator of which patients should receive supplemental services. Additional multicenter studies that evaluate a more comprehensive set of health outcomes related to pharmacist interventions in LTC are needed.

ACKNOWLEDGMENTS

We acknowledge the cooperation and nonfinancial support of the CCNC (a not-for-profit organization) as well as the physicians and pharmacists serving long-term facility residents, nursing home administrators, and the North Carolina Division of Medical Assistance. We particularly acknowledge the data access and computer support provided by AccessCare, a member of the CCNC. AccessCare also provided limited financial support for the production of this manuscript.

The authors thank the Long-Term Care Polypharmacy Alliance of North Carolina; the North Carolina Office of Research, Demonstrations and Rural Health Development; North Carolina Medical Directors Association; the North Carolina Health Care Facilities Association; and CCNC for their participation, program development, and assistance with collection and analysis of data. The authors also thank Coralynn B. Trewet, PharmD, University of Iowa; Brandy Newman, PharmD, King Soopers Pharmacy; and numerous pharmacy students from the University of North Carolina, Chapel Hill, for their assistance with data compilation and program development.

This paper was derived from a doctoral dissertation by Dr. Trygstad. The authors have indicated that they have no conflicts of interest regarding the content of this article.

REFERENCES

1. Beers MH, Ouslander JG, Rollingher I, et al, for the UCLA Division of Geriatric Medicine. Explicit criteria for determining inappropriate medication use in nursing home residents. *Arch Intern Med*. 1991;151:1825–1832.
2. Dhall J, Larrat EP, Lapane KL. Use of potentially inappropriate drugs in nursing homes. *Pharmacotherapy*. 2002;22: 88–96.
3. Hanlon JT, Schmader KE, Ruby CM, Weinberger M. Suboptimal prescribing in older inpatients and outpatients. *J Am Geriatr Soc*. 2001;49:200–209.
4. Grymonpre RE, Mitenko PA, Sitar DS, et al. Drug-associated hospital admissions in older medical patients. *J Am Geriatr Soc*. 1988;36:1092–1098.
5. Karch FE, Lasagna L. Adverse drug reactions. A critical review. *JAMA*. 1975;234:1236–1241.
6. Beers MH. Explicit criteria for determining potentially inappropriate medication use by the elderly. An update. *Arch Intern Med*. 1997;157:1531–1536.
7. Bero LA, Lipton HL, Bird JA. Characterization of geriatric drug-related hospital readmissions. *Med Care*. 1991;29: 989–1003.
8. Col N, Fanale JE, Kronholm P. The role of medication noncompliance and adverse drug reactions in hospitalizations of the elderly. *Arch Intern Med*. 1990;150:841–845.
9. Hanlon JT, Artz MB. Drug-related problems and pharmaceutical care: What are they, do they matter, and what's next? *Med Care*. 2001;39:109–112.
10. Adverse drug reactions. In: Delafuente JC, Stewart RB, eds. *Therapeutics in the Elderly*. 3rd ed. Cincinnati, Ohio: Harvey Whitney Books; 2001.
11. Field TS, Gurwitz JH, Avorn J, et al. Risk factors for adverse drug events among nursing home residents. *Arch Intern Med*. 2001;161:1629–1634.
12. NC Dept of Health and Human Services (NC DMA). Medicaid Statistics. NC Medicaid Annual Reports. http://www.dhhs.state.nc.us/dma/ncms.htm. Accessed December 15, 2008.
13. NC Dept of Health and Human Services (NC DMA). Table 11: North Carolina Medicaid State Fiscal Year 2003 Service Expenditures for Selected Major Medical Services by Program Category. http://www.dhhs.state.nc.us/dma/2003report/2003tables.pdf. Accessed December 15, 2008.
14. Odom CH, Wade T. The Department of Health and Human Services. Medicaid pharmacy management strategy—2003-2005. *N C Med J*. 2003;64:278–279.
15. Christensen D, Trygstad T, Sullivan R, et al. A pharmacy management intervention for optimizing drug therapy for nursing home patients. *Am J Geriatr Pharmacother*. 2004;2: 248–256.
16. Trygstad TK, Christensen D, Garmise J, et al. Pharmacist response to alerts generated from Medicaid pharmacy claims in a long-term care setting: Results from the North Carolina polypharmacy initiative. *J Manag Care Pharm*. 2005;11:575–583.
17. Beers MH. Explicit criteria for determining potentially inappropriate medication use by the elderly. An update. *Arch Intern Med*. 1997;157:1531–1536.
18. Fick DM, Cooper JW, Wade WE, et al. Updating the Beers criteria for potentially inappropriate medication use in

older adults: Results of a US consensus panel of experts [published correction appears in *Arch Intern Med*. 2004;164:298]. *Arch Intern Med*. 2003;163:2716–2724.
19. Weitzen S, Lapane KL, Toledano AY, et al. Principles for modeling propensity scores in medical research: A systematic literature review. *Pharmacoepidemiol Drug Saf*. 2004;13:841–853.
20. Perkins SM, Tu W, Underhill MG, et al. The use of propensity scores in pharmacoepidemiologic research. *Pharmacoepidemiol Drug Saf*. 2000;9: 93–101.
21. Rubin DB, Thomas N. Matching using estimated propensity scores: Relating theory to practice. *Biometrics*. 1996;52:249–264.
22. Wang J, Donnan PT. Propensity score methods in drug safety studies: Practice, strengths and limitations. *Pharmacoepidemiol Drug Saf*. 2001;10:341–344.
23. Drake C. Effects of misspecification of the propensity score on estimators of treatment effect. *Biometrics*. 1993;49:1231–1236.
24. Iezzoni LI, ed. *Risk Adjustment for Measuring Healthcare Outcomes*. 3rd ed. Chicago, Ill: Health Administration Press; 2003.
25. Rosenbaum PR. The consequences of adjustment for a concomitant variable that has been affected by the treatment. *J Royal Stat Soc*. 1984; 147:656–666.
26. Baser O. Too much ado about propensity score models? Comparing methods of propensity score matching. *Value Health*. 2006;9:377–385.
27. Dehejia RH, Wahba S. Propensity score matching methods for nonexperimental causal studies. *Rev Econ Stat*. 2002;84:151–161.
28. Leuven E, Sianesi B. PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. http://ideas.repec.org/c/boc/bocode/s432001.html. Accessed September 4, 2009.
29. Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *Am Stat*. 1985;39:33–38.
30. Trygstad TK, Garmise J, Christensen D. The North Carolina Polypharmacy Initiative. Report prepared for Nancy Henley, Deputy Director for Clinical Affairs, North Carolina Division of Medical Assistance. June 9, 2005.
31. Joffe MM, Rosenbaum PR. Invited commentary: Propensity scores. *Am J Epidemiol*. 1999;150:327–333.
32. Linden A, Adams JL, Roberts N. Using propensity scores to construct comparable control groups for disease management program evaluation. *Dis Manag Health Outcomes*. 2005;13:107–115.
33. Shadish WR, Cook TD, Campbell DT. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, Mass: Houghton Mifflin; 2001.
34. Baser O. Propensity score matching: Which method to choose? Web Presentation. Thomson Medstat. October 26, 2005.
35. Johnson A, Berg G, Fleegler E, Sauerbrun M. A matched-cohort study of selected clinical and utilization outcomes for an asthma care support program. *Dis Manag*. 2005;8: 144–154.
36. Berg GD, Wadhwa S, Johnson AE. A matched-cohort study of health services utilization and financial outcomes for a heart failure disease-management program in elderly patients. *J Am Geriatr Soc*. 2004;52:1655–1661.
37. Lapane KL, Hughes CM. Pharmacotherapy interventions undertaken by pharmacists in the Fleetwood phase III study: The role of process control. *Ann Pharmacother*. 2006;40: 1522–1526.
38. Adams AS, Zhang F, LeCates RF, et al. Prior authorization for antidepressants in Medicaid: Effects among disabled dual enrollees. *Arch Intern Med*. 2009;169:750–756.
39. Simoni-Wastila L, Ross-Degnan D, Mah C, et al. A retrospective data analysis of the impact of the New York triplicate prescription program on benzodiazepine use in Medicaid patients with chronic psychiatric and neurologic disorders. *Clin Ther*. 2004;26:322–336.
40. Soumerai SB. Benefits and risks of increasing restrictions on access to costly drugs in Medicaid. *Health Aff (Millwood)*. 2004;23:135–146.
41. Pearson SA, Ross-Degnan D, Payson A, Soumerai SB. Changing medication use in managed care: A critical review of the available evidence. *Am J Manag Care*. 2003;9:715–731.
42. Fortess EE, Soumerai SB, McLaughlin TJ, Ross-Degnan D. Utilization of essential medications by vulnerable older people after a drug benefit cap: Importance of mental disorders, chronic pain, and practice setting. *J Am Geriatr Soc.* 2001;49: 793–797.
43. Soumerai SB, Lipton HL. Computer-based drug-utilization review—risk, benefit, or boondoggle? *N Engl J Med*. 1995;332:1641–1645.
44. Soumerai SB, Ross-Degnan D, Avorn J, et al. Effects of Medicaid drug-payment limits on admission to hospitals and nursing homes. *N Engl J Med*. 1991;325:1072–1077.

Address correspondence to: Troy K. Trygstad, PharmD, MBA, PhD, 106 Kenilworth Place, Chapel Hill, NC 27516. E-mail: troy@t2email.com