

Clinical and economic outcomes in an observational study of COPD maintenance therapies: multivariable regression versus propensity score matching

Melissa H Roberts¹
Anand A Dalal²

¹Lovelace Clinic Foundation,
(Lovelace Respiratory Research
Institute at the time of the study),
Albuquerque, NM, ²US Health
Outcomes, GlaxoSmithKline,
Durham, NC, USA

Purpose: To investigate equivalency of results from multivariable regression (MR) and propensity score matching (PSM) models, observational research methods used to mitigate bias stemming from non-randomization (and consequently unbalanced groups at baseline), using, as an example, a large study of chronic obstructive pulmonary disease (COPD) initial maintenance therapy.

Methods: Patients were 32,338 health plan members, age ≥ 40 years, with COPD initially treated with fluticasone propionate/salmeterol combination (FSC), tiotropium (TIO), or ipratropium (IPR) alone or in combination with albuterol. Using MR and PSM methods, the proportion of patients with COPD-related health care utilization, mean costs, odds ratios (ORs), and incidence rate ratios (IRRs) for utilization events were calculated for the 12 months following therapy initiation.

Results: Of 12,595 FSC, 9126 TIO, and 10,617 IPR patients meeting MR inclusion criteria, 89.1% (8135) of TIO and 80.2% (8514) of IPR patients were matched to FSC patients for the PSM analysis. Methods produced substantially similar findings for mean cost comparisons, ORs, and IRRs for most utilization events. In contrast to MR, for TIO compared to FSC, PSM did not produce statistically significant ORs for hospitalization or outpatient visit with antibiotic or significant IRRs for hospitalization or outpatient visit with oral corticosteroid. As in the MR analysis, compared to FSC, ORs and IRRs for all other utilization events, as well as mean costs, were less favorable for IPR and TIO.

Conclusion: In this example of an observational study of maintenance therapy for COPD, more than 80% of the original treatment groups used in the MR analysis were matched to comparison treatment groups for the PSM analysis. While some sample size was lost in the PSM analysis, results from both methods were similar in direction and statistical significance, suggesting that MR and PSM were equivalent methods for mitigating bias.

Keywords: multivariate analysis, outcomes research, propensity score, pulmonary disease, chronic obstructive, statistical bias, statistical models

Introduction

Although considered the “gold standard” for evaluating treatment effectiveness, randomized clinical trials (RCTs) have important limitations. Because randomization removes potential bias from unknown and unmeasured confounders, observed differences in measured outcomes can be reasonably attributed to the treatment alone.¹ For valid experimental reasons, however, RCTs frequently restrict enrollment based on existing comorbidities, treatment history, and disease severity, among other criteria. As a result, outcomes observed in RCTs cannot necessarily be generalized to the

Correspondence: Melissa H Roberts
LCF Research, 2309 Renard Place SE,
Suite 103, Albuquerque, NM 87106,
USA
Tel +1 505 938 9922
Fax +1 505 938 9940
Email melissa.roberts@lcfresearch.org

“real world” of clinical practice, where patients present with varying degrees of disease severity and a range of comorbidity profiles. While there has been a call for increasing the number of pragmatic clinical trials, trials that examine outcomes among diverse populations of patients in typical practice settings are still rare.²

Retrospective, observational studies are valuable because they contribute pragmatic knowledge about treatment risk, effectiveness, and cost in clinical practice settings – knowledge that is critical to health care decision makers. In addition, observational studies are less costly and more quickly accomplished than RCTs, and can utilize large databases, permitting analysis of infrequent events. Because observational studies don't involve randomization of patients to treatment groups, however, selection bias can occur, and unmeasured variables can confound the associations between treatments and outcomes.

Multivariable regression (MR) methods are commonly used to control for confounding factors in observational studies. Matching is an alternative strategy. It can be done at the individual level, as in case–control matching, or at the group or frequency level, as in stratified random sampling. The matching process involves diagnostic checks regarding the balance of covariates across groups and provides information about the quality of the inferences that can be drawn from the subsequent analysis.³ Propensity score matching (PSM) has been increasingly used in epidemiologic studies of medical treatment effectiveness.¹ A propensity score represents the propensity of a particular subject to receive a particular treatment, based on the subject's pre-treatment characteristics.^{1,4,5} The score combines many covariates into a single variable and enables individuals from each treatment group with similar covariate values to be matched, as a quasi-randomization method.³ Subjects who cannot be matched are excluded from the analysis. An advantage of PSM is that matched sets with comparable covariate distributions can be created without the need for exact matches of each variable, which is problematic when there are more than a few covariates.³ Propensity matching works best if there is a fairly large overlap between the groups in terms of propensity to be given a treatment. When there is not, underlying selection bias may exist.³ Despite this method's theoretical benefits, in studies where both MR and PSM analysis methods have been used, only a small percentage of results (10% in one review and 13% in another) have been markedly different.^{1,6}

Disease exacerbations are important events in the course of COPD. Moderate and severe exacerbations adversely affect lung function, potentially accelerating disease progression.^{7,8}

Frequency of exacerbations is a significant factor in deteriorating health.^{9,10} Exacerbations also contribute substantially to health care utilization and costs. In the United States in 2010, the total cost of COPD was estimated to be US\$49.9 billion. Direct medical costs were estimated to be \$29.5 billion, including \$13.2 billion for COPD-related hospital care.¹¹ Reducing exacerbations is thus a singularly important goal of COPD management, both to improve patient quality of life and to reduce the indirect and direct medical costs of the disease. As pharmacotherapy is a primary means for reducing exacerbations, data concerning “real world” treatment effectiveness is of interest to health care providers, health care organizations, and health plans.

Agents for the relief and prevention of symptoms in COPD include short- and long-acting beta-agonists (including albuterol and salmeterol), short- and long-acting anticholinergics (including ipratropium bromide [IPR], tiotropium [TIO]), and inhaled corticosteroids (ICS).¹² Fluticasone propionate 250 µg/salmeterol 50 µg combination therapy (FSC) is an ICS plus long-acting beta-agonist used for treatment of airflow obstruction and reduction of exacerbations. Previously, we reported cost and utilization outcomes following initiation of COPD maintenance therapy with TIO, IPR (with or without albuterol), or FSC, using MR as the analysis method.¹³ To our knowledge, this was the first observational study to directly compare these three maintenance therapies. Compared to TIO and IPR, FSC was associated with lower COPD-related costs and utilization (hospitalizations, emergency department [ED] visits, and outpatient visits associated with an antibiotic or oral corticosteroid prescription). The objective of the present study was to perform a parallel analysis employing PSM to investigate the equivalency of results with those obtained in the prior MR analysis.

Materials and methods

Using PSM methods, we conducted a parallel analysis of COPD-related health care utilization and costs in patients with COPD receiving initial maintenance therapy (IMT) with FSC, TIO, or IPR, and we compared the results to those of a previous MR analysis. The term IMT refers to the patient's first instance of a pharmacy claim for a COPD maintenance medication; prior to this point, the patient's records indicated that he/she had not received maintenance therapy, only reliever medications or no medication.

We assessed exacerbations using claims data to measure health care utilization events related to exacerbations. There is no universally accepted definition of exacerbation. In clinical research, exacerbations generally are

defined in terms of worsening symptoms, unscheduled medical attention, and courses of antibiotics and/or oral corticosteroids.¹⁴ In observational studies such as ours, in which clinical and laboratory data are absent, exacerbations typically have been defined in terms of COPD-related health care utilization events, including hospitalizations, ED visits, physician visits, and outpatient pharmacy fills for oral corticosteroids/antibiotics.

Data source

Administrative data were obtained from the IMS LifeLink Health Plan Claims Database (IMS Health, Watertown, WA), which contains enrollment and demographic data, and health care and outpatient pharmacy claims from more than 40 million members of more than 70 US health plans. Calculated costs were based on allowed amounts, which most closely resemble the direct health care cost burden of illness. This is typically the amount the health plan pays, plus any member liability (eg, co-payment, deductible, or coinsurance amount). For claims with missing charges due to capitation arrangements, charges are imputed by IMS. The dataset included patient demographic and enrollment data, outpatient pharmacy claims, and medical services claims (outpatient, ED, and inpatient claims, including both facility claims and professional services claims) for January, 2004 to June, 2009. The specific content of the dataset has been described previously.¹³

Multivariable regression analysis

In the prior retrospective, observational cohort study, COPD-related clinical and economic outcomes were evaluated in patients who received one of three IMT medications for COPD (FSC, TIO, or IPR).¹³ The study perspective was that of the health plan provider organization, and only direct costs were considered. The study population included health plan members with diagnosed COPD who were new to maintenance therapy with FSC 250 µg/50 µg, TIO, or IPR (alone or in fixed dose combination with albuterol). The members were age 40 years and older, had a primary or secondary diagnosis of COPD (at least one ED visit, one hospitalization, or two outpatient visits with a primary or secondary International Classification of Disease, 9th edition, Clinical Modification [ICD-9-CM] diagnosis code of 491.xx, 492.xx, or 496.xx), had an IMT pharmacy claim between July 1, 2004 and June 30, 2008 (the date of the first identified prescription claim was the “index date”), had at least 6 months of continuous health plan enrollment prior to the index date (the “baseline period”), and at least

12 months of continuous enrollment following the index date (the “follow-up period”). Patients were excluded if they had a prescription drug fill for a COPD maintenance medication (FSC, IPR, TIO, budesonide/formoterol, inhaled corticosteroid alone, or long-acting beta-agonist alone) during the baseline period, or a pharmacy fill for an alternate study IMT medication (FSC, TIO, or IPR) within 60 days of the index date. Patients were excluded if they had a primary or secondary diagnosis of respiratory tract cancer (larynx, trachea, or pleura [ICD-9-CM codes of 161, 161.x, 162, 163, 163.x, 231, 231.x]) during the baseline period. The patient eligibility criteria and selection process have been described in detail previously.¹³

The primary utilization outcomes were incidence and mean number of COPD-related outpatient visits, outpatient visits associated with an antibiotic prescription fill, outpatient visits associated with an oral corticosteroid fill, hospitalizations, ED visits, and hospitalization and/or ED visits (combined endpoint). Encounters with a primary diagnosis code of 491.xx, 492.xx, or 496 were defined as COPD-related. The primary cost outcomes were mean COPD-related medical services costs, outpatient pharmacy costs (COPD controller and relief medications, oral corticosteroids, and antibiotics), and total costs (the sum of the two). Medical services costs comprised inpatient, outpatient, and ED care (including facility charges and professional service fees). Costs were inflated to 2009 dollars on a monthly basis using the medical care portion of the US Consumer Price Index.¹⁵ As outcomes were evaluated over a 12-month follow-up period, no discounting was applied to events or costs.

Bivariate analyses were used to compare differences between treatment cohorts in health care utilization and cost outcomes for the 12-month follow-up period. Multivariable logistic regression was used to model the risk for any health care utilization event as an odds ratio (OR). Negative binomial and Poisson regression were used to calculate incidence rate ratios (IRRs). Because of the right-skewed nature of the cost distribution, a generalized linear model using a gamma distribution was used to estimate differences in treatment costs. Estimates of mean differences and 95% confidence intervals (CIs) were calculated from the predicted cost values. The multivariable models controlled for age, sex, treatment, comorbidities (including asthma and heart disease), and COPD-related health care utilization at baseline.

Propensity score matching analysis

Starting with the original patient sample identified for the MR analysis, we created matched cohorts for the PSM analysis.

The TIO patients and IPR patients were separately matched to FSC patients based on propensity score; that is, patients initiating therapy with TIO were matched to patients initiating with FSC, and patients initiating therapy with IPR were matched to patients initiating with FSC. The matched samples were created based on each patient's predicted probability (propensity) of assignment to the case treatment (TIO or IPR). The propensity to be a patient whose initial maintenance therapy was TIO (or alternatively, IPR) incorporated the following baseline factors in the logistic regression equation: sex, age category, geographic region, comorbidities, COPD-related health care utilization, non-COPD-related health care utilization, COPD medication use, and COPD-related medical services costs. The utilization factors were hospitalization count and binary variables for outpatient visit, outpatient visit associated with an oral corticosteroid fill, outpatient visit associated with an antibiotic fill, ED visit, and hospitalization and/or ED visit (combined endpoint). Medication use was included using binary variables for short-acting beta-agonist (SABAs), oral corticosteroid, oral antibiotic, leukotriene modifier, and methylxanthine. Asthma, heart disease, and other relevant comorbidities were included as covariates. The greedy match algorithm was used, which performs matching using as much information as possible through the "nearest available pair" (or "nearest-neighbor") matching method with a caliper component.^{16–18} Once a match is made, the greedy algorithm does not reconsider the match. Because no available matches could be identified for some patients in the original cohorts, the patient sample for the PSM analysis was a smaller subset of the sample used in the MR analysis.

The utilization and cost outcomes reviewed were the same as for the MR analysis. Bivariate analyses were used to compare differences in outcomes in the 12-month period following initiation of maintenance therapy for the FSC-TIO and FSC-IPR matched cohorts. Logistic regression was used to model the risk for any health care utilization event (OR), and negative binomial and Poisson regression were used to calculate IRRs. Mean cost differences and 95% CIs were assessed using least squares estimates from generalized linear models using a gamma distribution. Since the PSM treatment groups were already matched for baseline characteristics, and our interest was only in the treatment effect, the PSM regression models contained only a factor for case IMT (TIO or IPR), with FSC used as the reference medication.

For both the MR and PSM analyses, statistical tests were two-sided, with an α -level of 0.05 for statistical significance. Demographic, clinical, and health care utilization

characteristics were assessed as counts and percentages for categorical variables, and as standard measures (mean and SD) for continuous variables. Both unpaired and paired *t*-tests were used for determining significant differences in mean measures. The chi-square test and McNemar's test were used to test paired proportion differences because significance tests that do not consider the non-independence of matched data have been found to be more conservative than tests for paired comparisons.^{19,20} Adequacy of matching was assessed using *P* values for comparison tests and standardized percentage differences.^{4,21} All analyses were conducted using SAS software (v 9.2 for Windows; SAS Institute, Cary, NC).

Results

A total of 32,338 patients met patient selection criteria in the MR analysis: 12,595 FSC patients, 9126 TIO patients, and 10,617 IPR patients. For the PSM analysis, 89.1% (8135) of the TIO patients were matched to FSC patients (64.6% of the original FSC cohort) and 80.2% (8514) of the IPR patients were matched to similar FSC patients (67.6% of the original FSC cohort). Figure 1 depicts the overlap in propensity scores between the groups. There was a large degree of overlap in the populations and good balance was achieved between the matched groups.

Baseline characteristics after matching

Baseline demographic, clinical, and utilization characteristics of the cohorts after matching on propensity score are shown in Table 1 (TIO-FSC) and Table 2 (IPR-FSC), along with *P* values for unpaired significance tests. Paired significance tests showed similar results although the *P* values were almost always lower (data not shown). After matching, the TIO and FSC groups were well balanced with respect to baseline characteristics; the groups were different only in mean COPD-related outpatient visits ($P < 0.001$). Matching between the IPR and FSC patients involved more factors. After matching, differences were present for some baseline characteristics: mean COPD-related outpatient visits ($P = 0.02$), mean all-cause outpatient visits ($P < 0.001$), and mean days' supply of SABAs ($P = 0.007$). Figures 2 and 3 show absolute standardized difference percentages for baseline characteristics prior to and after matching. These graphs further illustrate that, while some significant differences remained after matching, they were small from a clinical standpoint. Absolute standardized percentage differences were less than 10% for all assessed baseline characteristics in both groups, supporting an assessment of balance between groups.²²

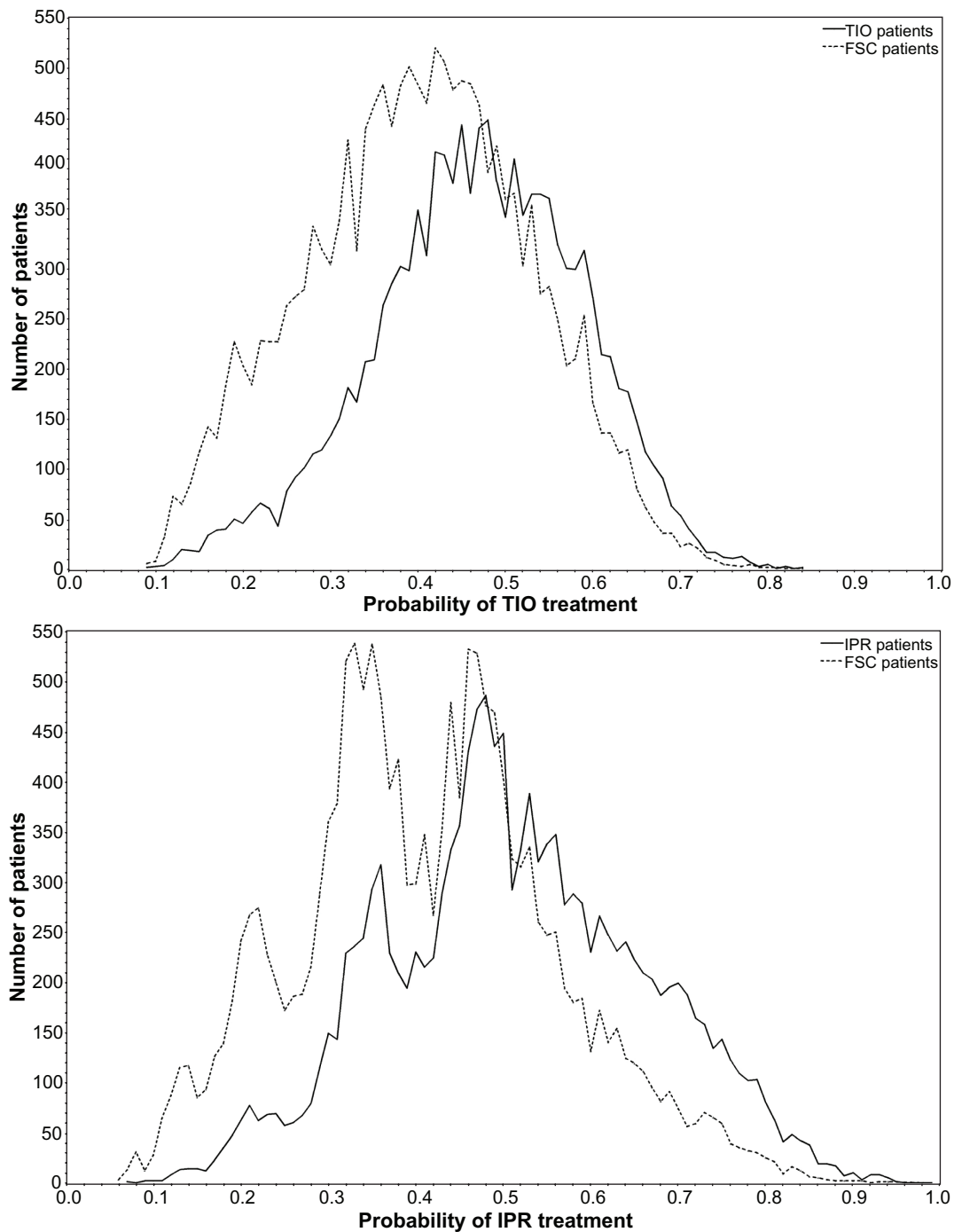


Figure 1 Overlap between FSC and TIO (top graph) and FSC and IPR (bottom graph) groups for calculated probability to receive TIO or IPR treatment.

Notes: Probability was calculated using logistic regression and adjusted for baseline covariates. The large intersection of the TIO and FSC distributions (89.1% of TIO patients matched) and IPR and FSC distributions (80.2% of IPR patients matched) indicates a high degree of similarity on the measured attributes.

Abbreviations: FSC, fluticasone propionate/salmeterol combination; IPR, ipratropium; TIO, tiotropium.

We compared differences in baseline characteristics between excluded patients and matched patients (Figures 2 and 3). Characteristics with large differences between excluded and matched patients tended to be the same characteristics as those associated with large standardized differences prior to matching. The excluded TIO and IPR patients were older, and

had more comorbidities and higher health care utilization. The excluded TIO patients, when compared to FSC patients who were not matched to TIO patients, were older (mean, 66.9 vs 60.1 years, $P < 0.001$) and more likely to be male (68.7% vs 39.8%, $P < 0.001$), not to have asthma (4.4% vs 46.8%, $P < 0.001$), to have lower use of leukotriene modifiers

Table 1 Demographic and clinical characteristics of tiotropium and fluticasone propionate/salmeterol cohorts matched on propensity score

	FSC n = 8135	TIO n = 8135	P value ^a
Age, mean (SD)	64.3 (11.57)	64.2 (11.37)	0.48
Male, n (%)	4173 (51.3)	4187 (51.5)	0.83
Comorbid conditions, n (%)			
Alcohol abuse	116 (1.4)	130 (1.6)	0.37
Anemia	87 (1.1)	90 (1.1)	0.82
Arrhythmia	1103 (13.6)	1111 (13.7)	0.85
Asthma	1017 (12.5)	1071 (13.2)	0.21
Congestive heart failure	999 (12.3)	1018 (12.5)	0.65
Diabetes (uncomplicated)	1554 (19.1)	1554 (19.1)	1.00
Fluid disorders	561 (6.9)	605 (7.4)	0.18
Heart disease	2020 (24.8)	2047 (25.2)	0.62
Hypertension (uncomplicated)	4010 (49.3)	4054 (49.8)	0.49
Hypothyroidism	692 (8.5)	667 (8.2)	0.48
Obstructive sleep apnea	624 (7.7)	629 (7.7)	0.88
Other lung conditions	1698 (20.9)	1750 (21.5)	0.32
Other neurological disease	255 (3.1)	237 (2.9)	0.41
Renal failure	335 (4.1)	335 (4.1)	1.00
Valvular disease	840 (10.3)	845 (10.4)	0.90
Rescue medicine use, n (%)			
SABA use	1677 (20.6)	1662 (20.4)	0.77
COPD-related utilization, number of encounters: mean (SD)			
Outpatient visit	0.58 (1.13)	0.64 (1.25)	0.001
Outpatient visit with antibiotic	0.06 (0.28)	0.06 (0.29)	1.00
Outpatient visit with oral corticosteroid	0.03 (0.20)	0.03 (0.24)	0.75
ED visit	0.02 (0.15)	0.02 (0.16)	0.58
Hospitalization	0.05 (0.24)	0.06 (0.24)	0.60
Hospitalization or ED visit	0.08 (0.29)	0.08 (0.29)	0.46
COPD-related costs (\$US), mean (SD)			
Medical services ^b	958 (6346)	982 (5780)	0.81
Pharmacy	34 (113)	35 (112)	0.40
Total	992 (6346)	1017 (5781)	0.80

Notes: ^aCalculated using chi-square test for frequencies and t-test for continuous measures, P values for cost measures are calculated using the Wilcoxon rank-sum test; ^bhealth care facility charges and professional services fees.

Abbreviations: COPD, chronic obstructive pulmonary disease; ED, emergency department; FSC, fluticasone propionate/salmeterol combination; SABA, short-acting beta-agonist; TIO, tiotropium.

(2.4% vs 12.6%, $P < 0.001$), and SABAs (15.4% vs 32.9%, $P < 0.001$), and to have significantly higher COPD-related medical service costs (US\$3734 vs \$365, $P < 0.001$). Similarly, excluded IPR patients, when compared to FSC patients not matched to IPR patients, were older (68.6 vs 60.4 years, $P < 0.001$) and more likely to be male (58.4% vs 37.0%, $P < 0.001$), not to have asthma (6.6% vs 45.3%, $P < 0.001$), to have lower use of leukotriene modifiers (1.5% vs 14.6%, $P < 0.001$) and SABAs (6.2% vs 45.1%, $P < 0.001$), and to have significantly higher COPD-related medical service costs (\$5437 vs \$220, $P < 0.001$).

Table 2 Demographic and clinical characteristics of ipratropium and fluticasone propionate/salmeterol cohorts matched on propensity score

	FSC n = 8514	IPR n = 8514	P value ^a
Age, mean (SD)	64.0 (12.03)	64.2 (12.15)	0.29
Male, n (%)	4173 (49.0)	4090 (48.0)	0.20
Comorbidities, patients with diagnosis: n (%)			
Alcohol abuse	142 (1.7)	139 (1.6)	0.86
Anemia	85 (1.0)	93 (1.1)	0.55
Arrhythmia	1173 (13.8)	1192 (14.0)	0.67
Asthma	1269 (14.9)	1324 (15.6)	0.24
Congestive heart failure	1132 (13.3)	1156 (13.6)	0.59
Depression	615 (7.2)	620 (7.3)	0.88
Diabetes (uncomplicated)	1854 (21.8)	1865 (21.9)	0.84
Fluid disorders	729 (8.6)	746 (8.8)	0.64
Heart disease	2060 (24.2)	2099 (24.7)	0.49
Hypertension (uncomplicated)	4191 (49.2)	4230 (49.7)	0.55
Hypothyroidism	679 (8.0)	701 (8.2)	0.54
Obstructive sleep apnea	601 (7.1)	560 (6.6)	0.21
Other lung conditions	1730 (20.3)	1800 (21.1)	0.19
Other neurological disease	299 (3.5)	302 (3.5)	0.90
Pulmonary circulation	206 (2.4)	212 (2.5)	0.77
Renal failure	393 (4.6)	395 (4.6)	0.94
Valvular disease	864 (10.1)	878 (10.3)	0.72
Weight loss	50 (0.6)	72 (0.8)	0.05
Rescue medicine use, n (%)			
SABA use	1296 (15.2)	1320 (15.5)	0.61
COPD-related utilization, number of encounters: mean (SD)			
Outpatient visit	0.48 (1.08)	0.52 (1.22)	0.02
Outpatient visit with antibiotic	0.07 (0.29)	0.07 (0.30)	0.23
Outpatient visit with oral corticosteroid	0.04 (0.24)	0.04 (0.23)	0.87
ED visit	0.03 (0.18)	0.03 (0.18)	0.93
Hospitalization	0.06 (0.25)	0.06 (0.26)	0.32
Hospitalization or ED visit	0.09 (0.31)	0.09 (0.31)	0.44
COPD-related costs (\$US), mean (SD)			
Medical services ^b	998 (6318)	1080 (6692)	0.42
Pharmacy	31 (106)	34 (138)	0.11
Total	1029 (6319)	1114 (6692)	0.40

Notes: ^aCalculated using chi-square test for frequencies and t-test for continuous measures, P values for cost measures are calculated using the Wilcoxon rank-sum test; ^bhealth care facility charges and professional services fees.

Abbreviations: COPD, chronic obstructive pulmonary disease; ED, emergency department; FSC, fluticasone propionate/salmeterol combination; IPR, ipratropium; SABA, short-acting beta-agonist.

Health care utilization and costs

Utilization and cost outcomes in the 12 months following initiation of maintenance therapy are shown in Table 3 and Figure 4, respectively. For the utilization comparisons, the more conservative unpaired significance tests were used. As described above, some subjects in the original MR cohorts were excluded from the PSM cohorts during the matching process. Those excluded were predominantly

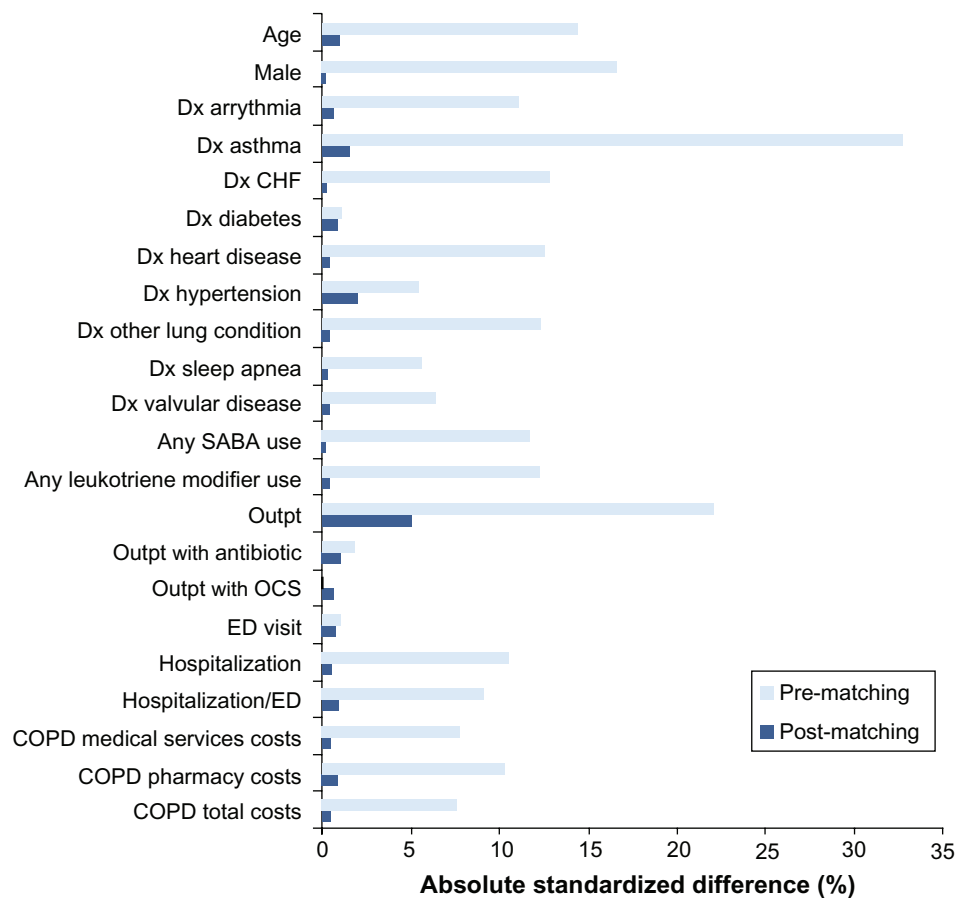


Figure 2 Standardized differences (percentages) in baseline characteristics between tiotropium and fluticasone propionate/salmeterol groups before and after matching. **Abbreviations:** CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; Dx, diagnosis; ED, emergency department; OCS, oral corticosteroid; Outpt, outpatient; SABA, short-acting beta-agonist.

older (and costlier) TIO and IPR patients, and younger FSC patients. This resulted in changes in the frequencies and means for outcomes in all treatment groups in the PSM analysis. However, the shifts for utilization parameters tended to be small. For example, in the MR analysis, 3.6% of FSC patients, 4.7% of TIO patients, and 7.3% of IPR patients had one or more ED visit ($P < 0.001$ for all differences between TIO and FSC and between IPR and FSC).¹⁹ In the TIO-FSC PSM analysis, 3.4% of FSC patients and 4.5% of TIO patients had one or more ED visit. In the IPR-FSC PSM analysis, 3.8% of FSC patients and 6.6% of IPR patients had one or more ED visit ($P < 0.001$ for both comparisons). Thus, in both analysis methods, the incidence of ED visits was lower in the FSC group, and differences between treatment groups were similar in magnitude. The excluded FSC patients had almost no impact on mean cost estimates for patients treated with FSC. However, for patients treated with TIO and IPR, the exclusion of older and sicker patients resulted in lower cost estimates for COPD-related medical services and total COPD-related costs (Figure 4).

TIO versus FSC

Several significant differences between the TIO and FSC groups seen in the MR analysis were also seen in the PSM analysis. The FSC group had a lower percentage of patients with an outpatient visit, outpatient visit associated with an oral corticosteroid, ED visit, or hospitalization/ED visit. In contrast to the MR analysis, the PSM analysis found no difference in the percentage of patients with a hospitalization ($P = 0.25$) or outpatient visit associated with an oral corticosteroid ($P = 0.08$). For each outcome measure, the percentage of patients with an encounter was lower in the FSC cohort than in the TIO and IPR cohorts, although, in the PSM analysis, because of the excluded younger FSC and older TIO patients, the FSC percentages increased slightly and the TIO percentages decreased slightly, diminishing the absolute differences between the two groups. With the exception of pharmacy costs, differences in costs that were significant in the MR analysis were also significant in the PSM analyses. FSC was associated with lower medical services costs (FSC, US\$1085 [95% CI: \$1061–1108]; TIO,

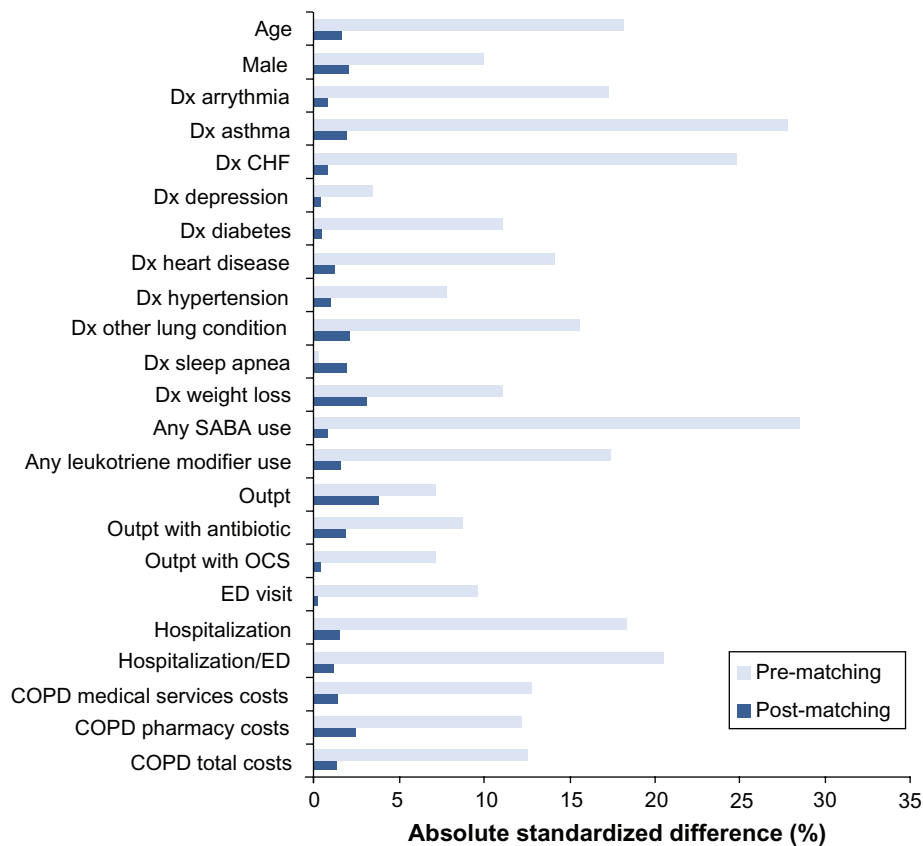


Figure 3 Standardized difference percentages in baseline characteristics between ipratropium and fluticasone propionate/salmeterol groups before and after matching.

Abbreviations: CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; Dx, diagnosis; ED, emergency department; OCS, oral corticosteroid; Outpt, outpatient; SABA, short-acting beta-agonist.

US\$1316 [95% CI: \$1288–1345]), and total health care costs compared to TIO (FSC, \$2037 [95% CI: \$1993–2081]; TIO, US\$2267 [95% CI: \$2218–2316]).

IPR versus FSC

The original MR analysis found that, in each of the five categories of utilization events, a lower percentage of FSC patients compared to IPR patients experienced events. These findings were essentially duplicated in the PSM analysis, despite the exclusion of 20% of the IPR patients. (*P* values for all differences were <0.001 in the MR analysis, and ranged from <0.001 to 0.03 in the PSM analysis). Differences in COPD-related costs that were significant in the MR analysis were also significant in the PSM analysis. FSC was associated with higher pharmacy costs (FSC, \$917 [95% CI: \$897–936]; IPR, US\$614 [95% CI: \$601–627]), but lower medical service costs (FSC, \$1122 [95% CI: \$1099–1146]; IPR, US\$1746 [95% CI: \$1709–1784]), and total costs compared to IPR (FSC, US\$2039 [95% CI: \$1996–2083]; IPR, US\$2360 [95% CI: \$2311–2411]).

Risk for health care utilization

The ORs for health care utilization are shown in Figure 5. The MR and PSM analyses produced fairly similar ORs for various categories of health care utilization, with ORs produced by the PSM analysis being slightly lower. For example, in the MR analysis, the statistically significant hospitalization/ED visit ORs for TIO and IPR (with respect to FSC) are 1.28 and 1.72, respectively; these values are 1.21 and 1.67 in the PSM analysis, respectively. Nonetheless, both analyses show that TIO and IPR patients have higher ORs, compared to FSC patients, for outpatient visit, outpatient visit with oral corticosteroid, ED visit, and hospitalization/ED visit. The IPR and FSC comparison also showed higher ORs for hospitalization and for an outpatient visit with an antibiotic. However, while the MR analysis calculated slightly higher odds for hospitalization for TIO (OR: 1.19 [95% CI: 1.04–1.37]) compared to FSC, the PSM analysis found no difference (OR: 1.10 [95% CI: 0.94–1.28]), nor was any difference in risk found between TIO and FSC for an outpatient visit with an antibiotic (OR: 1.14 [95% CI: 0.98–1.32]).

Table 3 Number of patients with COPD-related health care utilization in the 12 months following therapy initiation: comparison of TIO to FSC and IPR to FSC using two analysis methods

COPD-related utilization category	Multivariable regression analysis ¹³ patients with any encounter: n (%)			Propensity score-matched analysis, patients with any encounter: n (%)		
	FSC n = 12,595	TIO n = 9126	P value ^a	FSC n = 8135	TIO n = 8135	P value ^a
Outpatient visit	3615 (28.7)	3661 (40.1)	<0.001	2567 (31.6)	3147 (38.7)	<0.001
Outpatient visit with antibiotic	490 (3.9)	478 (5.2)	<0.001	355 (4.4)	402 (4.9)	0.08
Outpatient visit with oral corticosteroid	261 (2.1)	262 (2.9)	0.001	178 (2.2)	224 (2.8)	0.02
ED visit	450 (3.6)	427 (4.7)	<0.001	277 (3.4)	366 (4.5)	<0.001
Hospitalization	446 (3.5)	413 (4.5)	<0.001	314 (3.9)	343 (4.2)	0.25
Hospitalization/ED visit	819 (6.5)	764 (8.4)	<0.001	544 (6.7)	647 (8.0)	0.002
	FSC n = 12,595	IPR n = 10,617	P value ^a	FSC n = 8514	IPR n = 8514	P value ^a
Outpatient visit	3615 (28.7)	3788 (35.7)	<0.001	2501 (29.4)	2940 (34.5)	<0.001
Outpatient visit with antibiotic	490 (3.9)	617 (5.8)	<0.001	358 (4.2)	454 (5.3)	<0.001
Outpatient visit with oral corticosteroid	261 (2.1)	354 (3.3)	<0.001	189 (2.2)	269 (3.2)	<0.001
ED visit	450 (3.6)	778 (7.3)	<0.001	322 (3.8)	566 (6.6)	<0.001
Hospitalization	446 (3.5)	651 (6.1)	<0.001	328 (3.9)	475 (5.6)	<0.001
Hospitalization/ED visit	819 (6.5)	1284 (12.1)	<0.001	594 (7.0)	945 (11.1)	<0.001

Notes: ^aCalculated using chi-square test.

Abbreviations: COPD, chronic obstructive pulmonary disease; ED, emergency department; FSC, fluticasone propionate/salmeterol combination; IPR, ipratropium; TIO, tiotropium.

Incidence rate ratios

The IRRs for health care utilization events in the TIO and IPR groups with reference to the FSC group are shown in Figure 5. Again, both analytic methods yielded fairly similar IRRs, with the PSM analysis producing slightly lower IRRs for all categories of utilization. In all comparisons in the PSM analysis, as in the MR analysis, IPR patients were found to be at significantly higher risk for events, compared to FSC patients. For the TIO group compared to the FSC group, all IRRs in the MR analysis were significantly higher. However, in the PSM analysis, IRRs for outpatient visits with oral corticosteroid and for hospitalizations were no longer significant.

Discussion

In this analysis of data from an observational, retrospective cohort study of initial maintenance therapies for COPD, we demonstrated the similarity of results using two analytic approaches to observational research. Specifically, we compared results from a PSM analysis with those from a previously published, parallel MR analysis.¹³ We found that both methods yielded similar health care utilization and cost outcomes. General agreement between MR and PSM methods has been found in other studies. In a review of 177 comparative method studies, Sturmer concluded that substantial changes in treatment effects were seen when point

estimates were calculated with and without adjustment for covariates, but that the method of adjustment itself – MR or PSM – made little difference.¹

In PSM, a high degree of propensity score overlap after matching is desirable in terms of internal validity. When overlap is minimal, unmeasured confounding bias in treatment groups probably cannot be resolved using either MR or PSM techniques.¹ In the present PSM analysis, large proportions of patients in both the TIO and IPR cohorts (89.1% and 80.2%, respectively) were matched to FSC patients, and there was substantial overlap in propensity scores between groups. In other words, matching produced cohorts with similar baseline characteristics. In general, the few statistically significant differences remaining after matching were minor in terms of effect size and practical significance, and had small absolute standardized differences. Characteristics that would be expected to considerably skew utilization outcomes, such as comorbid cardiovascular disease, were not different between matched groups.

Outcomes differed slightly between the PSM and MR analyses. Some differences may be due to the smaller PSM sample size, since the excluded patients were a contributing explanatory factor for the lower PSM utilization and cost estimates. While the MR analysis was a population-based study, the PSM analysis, as a result of the matching process, excluded some younger individuals, with minimal

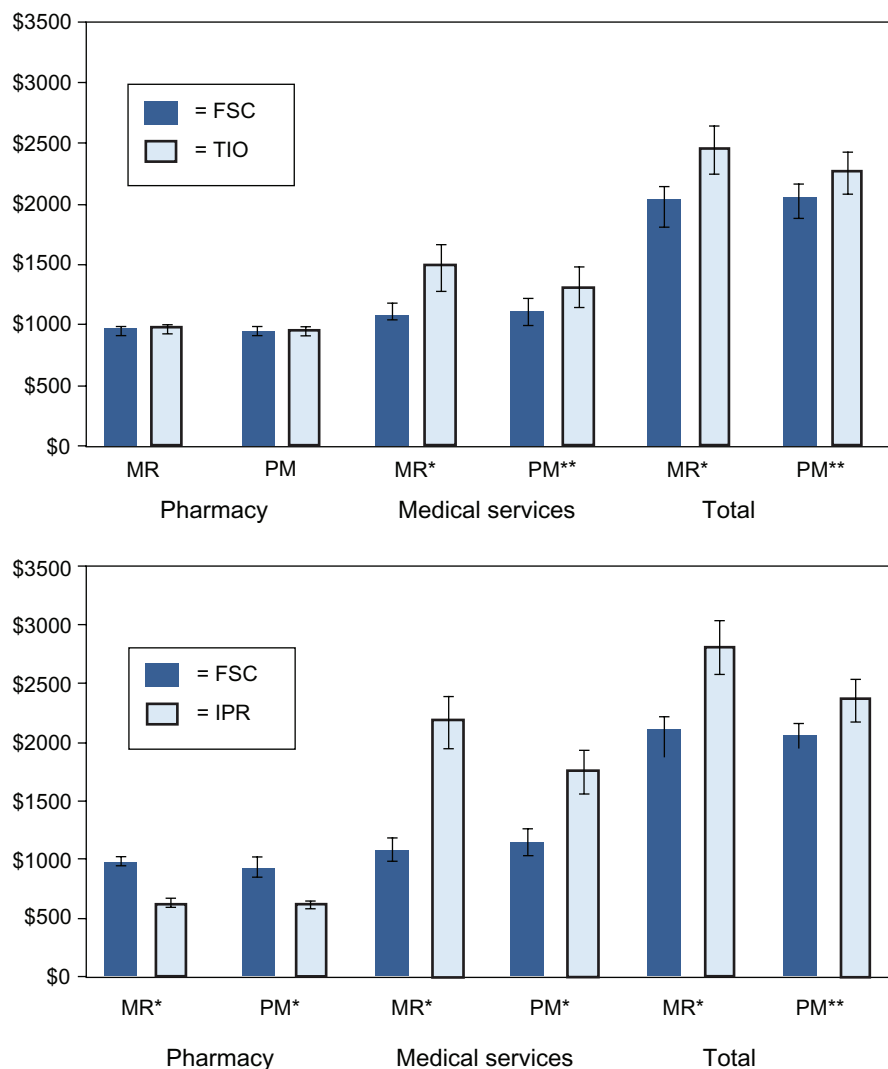


Figure 4 COPD-related health care costs in the 12 months following therapy initiation: comparison of TIO and IPR to FSC using two analysis methods.

Notes: Bars depict mean cost (2009 \$US) and 95% confidence interval. Medical services costs include health care facility charges and professional services fees; * $P < 0.001$; ** $P < 0.05$ (unpaired *t*-test).

Abbreviations: FSC, fluticasone propionate/salmeterol combination; IPR, ipratropium; MR, multivariable regression; PSM, propensity score matching; TIO, tiotropium.

comorbidities, who were treated with FSC, and some older, sicker individuals who were treated with IPR or TIO. Exclusion of the older and sicker individuals resulted in lower mean costs for IPR and TIO patients, while costs for FSC patients were quite similar in both analyses.

Event frequency also may be a factor in the differences in findings between analysis methods. Multivariable logistic regression and propensity matching have been found to produce similar results when events are not infrequent.^{23–25} Through simulations, Cepeda and colleagues found that the use of propensity scores yielded less biased estimates than multivariable logistic regression only when there were eight or fewer modeled events per covariate.²⁶ When the ratio of modeled events was higher, multivariable logistic

regression was the better method. Other studies have determined that ten events per covariate is a desired ratio when using maximum likelihood methods.²⁷ The main outcomes in our analyses, (COPD-related outpatient visits associated with an antibiotic or oral corticosteroid, ED visits, and hospitalizations), although of great concern clinically, occur relatively infrequently, from a statistical standpoint, when averaged across a large population of COPD patients that is unrestricted in terms of disease severity. Our original MR analysis included 44 covariates. The outcome of outpatient visit with an oral corticosteroid had the smallest number of events per covariate modeled, with 877 of 32,338 patients having at least one event, which translates to 20 events per covariate in the multivariate logistic regression analysis.

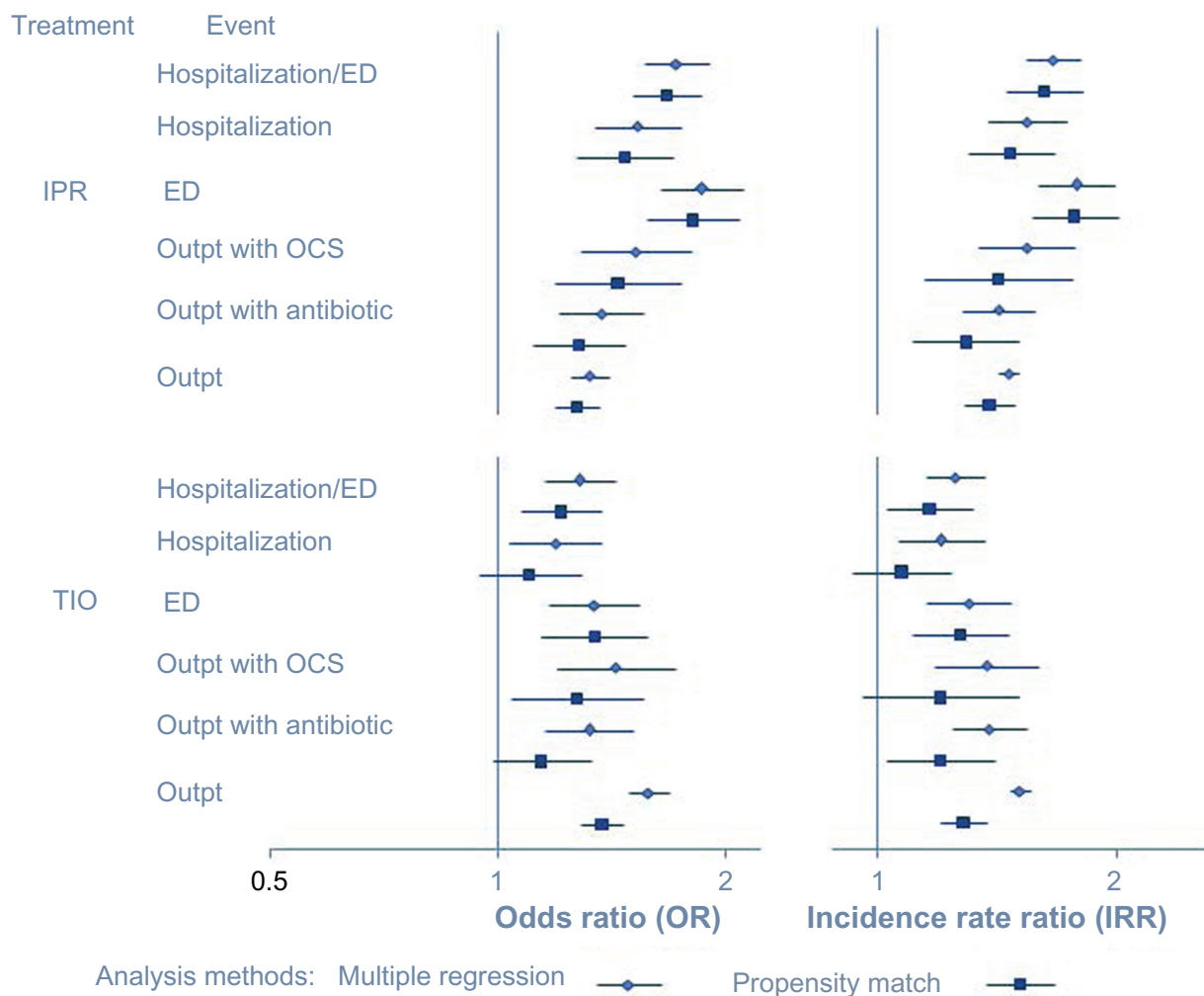


Figure 5 Risk (left graph) and incident rate ratios (right graph) for any health care encounter in the 12 months following therapy initiation: evaluation of TIO and IPR (with reference to FSC) using two analysis methods.

Abbreviations: ED, emergency department; FSC, fluticasone propionate/salmeterol combination; IPR, ipratropium; OCS, oral corticosteroids; Outpt, outpatient; TIO, tiotropium.

This compares to 65 events per covariate for the combined endpoint of hospitalization/ED visit. The lower ratios of events per covariate for some outcomes may have been a factor in the different findings of the two analyses for ORs and IRRs. On the other hand, costs (COPD-related medical service costs and pharmacy costs) were universally incurred, and both analyses found that, compared to FSC, TIO and IPR were associated with higher costs for COPD-related medical services, and higher total costs, even though costs associated with TIO and IPR were reduced in the PSM analyses.

Both MR and PSM methods adjust associations between treatment effects and outcomes to reduce potential bias from observed covariates. Other researchers have reported that results from the two methods appear to be consistent when there is large overlap between groups in propensity for a given treatment, which ensures minimal loss of observations,

and when outcomes can be modeled with a relatively large number of events per covariate.^{1,4,24} Our findings support this view and suggest that, with regard to less frequent events, in particular when effect sizes may be small, consideration should be given to analyzing outcomes using both methods, assuming a large proportion of subjects can be matched.

While PSM is a more transparent method, in the sense that it allows one to see the degree of equality between groups after matching, in this study, PSM provided little advantage over MR in terms of the validity of the results. Because of the inevitable reduction in sample size and change in overall composition of treatment groups being compared, the choice of whether to use PSM or MR will depend on the question being investigated, whether a population effect is being measured, and whether review of a non-representative population of patients receiving treatment is acceptable

(or even preferred). This point is addressed by D'Agostino, who recommended that when patients are excluded in matched analyses, researchers need to be particularly clear in their descriptions of the included and excluded patients, and of the populations to which study results are applicable.²⁸

Strengths of this study include the large sample sizes of both analyses and the high degree of propensity matching, with approximately 80% and 89% of the original IPR and TIO cohorts matched, respectively, to FSC patients. The exclusion of some original subjects from the PSM cohorts due to lack of a match does mean, however, that any additional information these subjects might have provided was lost, and statistical power affected. Some limitations should be considered in interpreting the results. We measured exacerbations using claims data, defining exacerbations as COPD-related health care events. Using an alternative definition of exacerbation based on symptoms, lung function, or other clinical parameters could influence observed effect sizes.¹⁴ However, we would not expect a different definition of exacerbations to influence effect sizes differently for MR than for PSM, or for it to change the overall findings of this study. Since both MR and propensity matched analyses attempt to reduce bias through adjustment using covariates, the ability to do this is dependent on the capture of all relevant factors. In this analysis, some potential confounders were unmeasured. As this was an observational study utilizing administrative claims data, information about patients' clinical status was not available. We could not ascertain lung function status, disease severity, or other clinical characteristics. Therefore, residual baseline differences between treatment groups may remain. However, we did control for two key characteristics of interest – disease severity and exacerbation frequency – by using prior COPD-related health care and pharmacy utilization (particularly oral corticosteroids/antibiotics) as proxy measures.

Conclusion

Results obtained in our analysis suggest that both MR and PSM methods are appropriate analytic techniques for addressing and mitigating bias in observational research. In this example of an observational study of maintenance therapy for COPD, more than 80% of the original treatment groups used in the MR analysis were matched to a comparison group for the PSM analysis. While some sample size was lost in the PSM analysis, results from both methods were similar in direction and statistical significance. Further, this analysis underscores the need for researchers to have a good understanding of the populations undergoing treatment and

the factors associated with both receipt of treatment and occurrence of the measured outcomes.

Acknowledgments

Melissa H Roberts contributed to the study concept and design, analysis and interpretation of data, drafting of the manuscript, statistical analysis, critical revision of the manuscript for important intellectual content, and final approval of the manuscript.

Anand A Dalal contributed to the study concept and design, interpretation of data, drafting of the manuscript, critical revision of the manuscript for important intellectual content, final approval of the manuscript, and administrative and material support.

The authors thank Christopher Blanchette and Hans Petersen (Lovelace Respiratory Research Institute, Albuquerque, NM) for data acquisition and management, and Judith S Hurley, (Hurley Health and Medical Communications, Albuquerque, NM) for medical writing services.

Disclosure

Melissa H Robert and Anand A Dalal had full access to all study data and take full responsibility for the integrity of the data and the accuracy of the data analysis. This study was sponsored by GlaxoSmithKline (Protocol number: ADC112646).

References

1. Sturmer T, Joshi M, Glynn RJ, Avorn J, Rothman KJ, Schneeweiss S. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. *J Clin Epidemiol*. 2006;59(5):437–447.
2. Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. *JAMA*. 2003;290(12):1624–1632.
3. Stuart EA. Matching methods for causal inference: a review and a look forward. *Stat Sci*. 2010;25(1):1–21.
4. D'Agostino RB. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med*. 1998;17(19):2265–2281.
5. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70(1):41–55.
6. Shah BR, Laupacis A, Hux JE, Austin PC. Propensity score methods gave similar results to regression modeling in observational studies: a systematic review. *J Clin Epidemiol*. 2005;58(6):550–559.
7. Seemungal TAR, Donaldson GC, Paul EA, Bestall JC, Jeffries DJ, Wedzicha JA. Effect of exacerbation on quality of life in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 1998;157(5):1418–1422.
8. Wedzicha JA, Seemungal TAR. COPD exacerbations: defining their cause and prevention. *Lancet*. 2007;370(9589):786–796.
9. Hurst JR, Vestbo J, Anzueto A, et al. Susceptibility to exacerbation in chronic obstructive pulmonary disease. *N Engl J Med*. 2010;363(12):1128–1138.

10. Spencer S, Calverley PMA, Burge PS, Jones PW. Impact of preventing exacerbations on deterioration of health status in COPD. *Eur Resp J*. 2004;23(5):698–702.
11. National Heart, Lung and Blood Institute. *Morbidity and mortality: 2009 chart book on cardiovascular, lung, and blood diseases*. Bethesda, MD: National Institutes of Health; 2009.
12. Global Initiative for Chronic Obstructive Lung Disease (GOLD). *Global Strategy for Diagnosis, Management and Prevention of COPD: Update 2010*. Global Initiative for Chronic Obstructive Lung Disease; 2010. Available from: <http://www.goldcopd.org>. Accessed on January 25, 2012.
13. Dalal AA, Roberts MH, Petersen HV, Blanchette CM, Mapel DW. Comparative cost-effectiveness of a fluticasone-propionate/salmeterol combination versus anticholinergics as initial maintenance therapy for chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis*. 2010;6:13–22.
14. Effing TW, Kerstjens HAM, Monninkhof EM, et al. Definitions of exacerbations: does it really matter in clinical trials on COPD? *Chest*. 2009;136(3):918–923.
15. Bureau of Labor Statistics, US Department of Labor. Measuring price change for medical care in the CPI. Available from: <http://www.bls.gov/cpi/cpifact4.htm>. Accessed January 25, 2012.
16. Davis G, Mallat S, Avellaneda M. Adaptive greedy approximations. *Constr Approx*. 1997;13(1):57–98.
17. Katajainen J, Raita T. An analysis of the longest match and the greedy heuristics in text encoding. *J Assoc Comput Machinery*. 1992;39(2):281–294.
18. Parsons LS. Reducing bias in a propensity score matched-pair sample using greedy matching techniques. *Proceedings of the Twenty-Sixth Annual SAS Users Group International Conference*; April 22–25, 2001; Long Beach, CA. Cary, NC: SAS Institute; 2001: 214–226
19. Austin PC. Type I error rates, coverage of confidence intervals, and variance estimation in propensity-score matched analyses. *Int J Biostat*. 2009;5(1): Article 13.
20. Li L. Comment: analyzing propensity score matched count data. *Int J Biostat*. 2010;6(1): Article 5.
21. Rosenbaum PR, Rubin DB. Constructing a control-group using multivariate matched sampling methods that incorporate the propensity score. *Am Stat*. 1985;39(1):33–38.
22. Normand ST, Landrum NB, Guadagnoli E, et al. Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores. *J Clin Epidemiol*. 2001;54(4):387–398.
23. Cadarette SM, Gagne JJ, Solomon DH, et al. Confounder summary scores when comparing the effects of multiple drug exposures. *Pharmacoepidemiol Drug Saf*. 2010;19(1):2–9.
24. Hemmila MR, Birkmeyer NJ, Arbabi S, et al. Introduction to propensity scores: a case study on the comparative effectiveness of laparoscopic vs open appendectomy. *Arch Surg*. 2010;145(10):939–945.
25. Spahn J, Sheth K, Yeh WS, et al. Dispensing of fluticasone propionate/salmeterol combination in the summer and asthma-related outcomes in the fall. *J Allergy Clin Immunol*. 2009;124(6):1197–1203.
26. Cepeda MS, Boston R, Farrar JT, et al. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. *Am J Epidemiol*. 2003;158(3):280–287.
27. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15(4):361–387.
28. D’Agostino RB. Discussion of: statistical and regulatory issues with the application of propensity score analysis to nonrandomized medical device clinical studies. *J Biopharm Stat*. 2007;17(1):29–33.

International Journal of COPD

Publish your work in this journal

The International Journal of COPD is an international, peer-reviewed journal of therapeutics and pharmacology focusing on concise rapid reporting of clinical studies and reviews in COPD. Special focus is given to the pathophysiological processes underlying the disease, intervention programs, patient focused education, and self management protocols.

Submit your manuscript here: <http://www.dovepress.com/international-journal-of-copd-journal>

Dovepress

This journal is indexed on PubMed Central, MedLine and CAS. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.