EFFECT OF ACADEMIC DETAILING ON COX-2 UTILIZATION RATES

By

STEPHEN DOUGLAS GRAHAM

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2005

To Evie

ACKNOWLEDGMENTS

I would like to thank my wife, Andrea, and sons, Nicolas and Andrew, for their love and support.

I thank my dissertation chair, Dr. Abraham Hartzema, and committee members, Drs. Ingrid Sketris, Almut Winterstein, Richard Segal, and Babette Brumback for their guidance through the dissertation process.

I would like to extend special thanks to Ms. Dawn Frail at the Nova Scotia Department of Health and again to Dr. Ingrid Sketris at Dalhousie University for providing me with overwhelming support and encouragement to succeed and to return to Canada.

Finally, I would like to thank the graduate students for giving me many happy memories of Florida.

TABLE OF CONTENTS

APPENDIX

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

EFFECT OF ACADEMIC DETAILING ON COX-2 UTILIZATION RATES

By

Stephen Douglas Graham

December, 2005

Chair: Abraham Hartzema
Major Department: Pharmacy Health Care Administration

Background: The prevalence of osteoarthritis (OA) is estimated at 50 to 80 % of the elderly population and therapy aims to relieve symptoms since there is no cure. Nova Scotia general practitioners (GPs) identified a need for an academic detailing (AD) intervention aimed at optimizing the management of OA.

Objectives: The primary objective was to measure the effect of an OA AD intervention to reduce the utilization rate of COX-2 inhibitors in the elderly population. Secondary objectives were to examine the intervention effect on the utilization rates of gastro-protective agents and medical services.

Methods: We conducted a retrospective cohort study employing administrative data to examine the effects of the intervention. Differences in utilization rates were evaluated using generalized estimating equation (GEE) analysis for longitudinal data.

Selection bias was anticipated since the intervention was voluntary, and randomization not possible. Three methods of propensity score (PS) analysis (quintile

stratification, regression on the PS, and "greedy matching") were evaluated for the ability to adjust for bias on PS model covariates.

Findings: We identified a significant difference in the change in COX-2 utilization rates between groups for the three month period following the intervention ($p = 0.0395$, 95% CI (0.0365, 1.4815)) and a significant decrease in the intervention group's within group utilization rate between the pre and post intervention periods ($z = -2.34$, $p = 0.0191$). The GP office visit rate was the only secondary outcome where the intervention group was significantly higher ($p = 0.0275$, 95% CI (-0.7926, -0.0464)). The difference occurred in the time period from three to six months post intervention.

Conclusions: The OA AD intervention was associated with a significant decrease in COX-2 utilization rates in the three month period immediately following the intervention. The effect of decreased utilization continued for the rest of the post intervention periods but was not statistically significant. The only secondary outcome to show a significant between groups effect was the GP office visit rate which was higher for the intervention group in the second three month post intervention time period.

CHAPTER 1
INTRODUCTION

Background

In June 2002, the Division of Continuing Medical Education (CME), Dalhousie University Faculty of Medicine, began their second academic detailing (AD) intervention with provincial physicians aimed at optimizing the care of osteoarthritis (OA) within the seniors population (persons greater than 65 years of age). The AD program is an ongoing initiative funded by the Nova Scotia Department of Health and managed by the Drug Evaluation Alliance of Nova Scotia (DEANS). As the AD program is a continuing effort and represents a significant cost to DEANS it is necessary to evaluate the effectiveness of the intervention.

The OA topic was chosen as an AD intervention based on the extent to which OA affects the elderly population and on the feedback that Dalhousie CME received from general practitioners (GPs) in a survey filled out following the previous influenza AD intervention which indicated the GPs' desire to have an OA AD intervention developed. The Dalhousie CME Division then presented the OA topic to a GP focus group where the need for education pertaining to available OA therapies was determined.

OA is a progressive disease that affects the joint cartilage and eventually leads to joint failure.[3] The prevalence of OA in the population is extremely high. It is estimated that 50 to 80% of the elderly population experience symptomatic OA.[4] Estimates specific to the province of Ontario, propose that almost all persons over the age of 65 exhibit signs of OA on radiographic evidence and of these 33% are symptomatic.[5] OA is equally

prevalent in men and women, with women showing more manifestation in the knees and hands and men more prevalent in the hip. Arthritis has been associated with half of all disability in the elderly population.[4] There is no known cure for OA[3] and available palliative treatments are associated with substantial toxicity and side effects.[4] Treatment is therefore primarily aimed at reducing pain, improving joint mobility, and limiting functional disability. Patient education regarding medications used in the treatment of OA (primarily for the control of pain) and appropriate exercise regimens is also important.[3, 4]

The OA AD intervention has set four learning objectives. Each physician visit will include at least the following: (1) a discussion of the goals of therapy, (2) recommendations for non-pharmacological treatments when appropriate (e.g., physio-therapy and exercise), (3) advice for patients about the safety and efficacy of acetaminophen, and (4) a discussion of the role of traditional non-steroidal anti-inflammatory drugs (NSAIDs).[6]

The primary interest of this research dealt with the fourth message specifically, the analysis of the effectiveness of the OA AD intervention as it pertains to the pharmacotherapy of OA and in particular the usage of COX-2 inhibitors. The Nova Scotia OA AD was developed in 2002 and the intervention called for the use of acetaminophen as a first line therapy for mild to moderate OA. The intervention suggested that if acetaminophen did not control pain symptoms, then the use of traditional NSAIDs in as low a dose as possible and for as short duration as possible was indicated. NSAIDs were considered appropriate therapy for moderate to severe OA.[6]

The role of COX-2 inhibitors in the management of OA was assessed by the OA AD group as controversial. The Ontario Treatment Guidelines for OA recommend that, based on evidence of similar efficacy and early evidence of somewhat lower rates of serious GI events, selective COX-2 inhibiting NSAIDs can be considered for patients at high risk of serious GI events.[3] This recommendation, however, is one that is from well-designed, randomized controlled trials or meta-analyses with inconsistent results or demonstrating equivocal benefit.[3] The Nova Scotia program states that, the precise role of COX-2 inhibitors in the treatment of OA remains to be determined.[6] The summary statements in the OA AD intervention[6] relay two points that are relevant to this analysis. Firstly, COX-2 inhibitors are as effective but not more effective than traditional NSAIDs for symptomatic treatment of OA and secondly, the CLASS[7] and VIGOR[8] trials were inconclusive in the analysis of the gastro-protective effects of COX-2 inhibitors.

When faced with the substantial effect of OA on the population,[6] the uncertain role of COX-2 inhibitors in the treatment of OA, the increased cost of COX-2 inhibitors over the traditional NSAIDs (appendix c), and the utilization rate of COX-2 inhibitors in the Nova Scotia pharmacare population of approximately 6% in 2001,[9] the DEANS Management Committee undertook to develop the AD intervention on OA Management.

Problem Statement

The effect of AD on clinical and economic outcomes is of great interest to the Nova Scotia government's policy makers as funding for interventions to improve the health care system is scarce. This research addresses the question of whether the AD program on OA is effective in lowering the utilization rate of COX-2 inhibitors. At the same time the study measures the effects that the program has on the utilization rates of other healthcare resources such as hospital or physician visits that occur as a result of GI side

effects associated with drug therapy with traditional NSAIDs and COX-2 inhibitors. The efficacy of traditional NSAIDs and COX-2 inhibitors in relieving pain is similar but the GI side effects profile for traditional NSAIDs is higher.[10] It is expected that the intervention could increase the utilization rates of gastro-protective agents (particularly misoprostol and proton pump inhibitors (PPIs)) but it is not expected to increase other health care utilization rates and will therefore not have negative impacts on the outcomes of care.

The methodological challenge for the evaluation of the OA AD intervention is the need to significantly adjust for selection bias that is likely present since GPs can choose to participate and those that do participate might be different from those that do not participate. Statistical adjustment through regression on the propensity score (PS) methods have been shown to be effective in reducing between group biases on many confounding variables.[11, 12] The use of PSs in studies that examine the unit of analysis other than the patient is uncommon in the medical literature. In this study the unit of measure was the GP. No other studies with the GP as the unit of measure were found in the medical literature so the evaluation of different PS method's ability to adjust for bias between GP groups was warranted.

Research Questions and Hypotheses

The term statistically significant is defined as results where the type I error (alpha) is less than 0.05. The results are statistically significant if the analysis yields p-values less than 0.05.

Hypotheses relating to research questions one to three are examining the effect of the OA AD intervention in the Nova Scotia residents who are greater than 65 years old

and have a GP who has participated in the intervention as compared with GPs in the province who did not participate in the intervention.

The first research question examined the expectation that GPs will consider the information provided in the OA AD intervention and choose not to prescribe COX-2 inhibitors for their elderly patients.

Research Question 1

Do the patients of GPs who have undertaken the OA AD intervention have significantly lower COX-2 inhibitor utilization rates after the GP has undergone the AD intervention as compared to a GP control cohort? (Are there significant between group differences?)

Research Question 1 Hypothesis

The null hypothesis is that the OA AD intervention will have no effect on the utilization rate of COX-2 inhibitors.

The alternative hypothesis is that the OA AD intervention will have the effect of decreasing the utilization rate of COX-2 inhibitors.

The second research question examined the sustainability of the intervention (if research question 1 hypothesis is found to be significant) since a shortcoming of the OA AD intervention (appendix a) is the lack of a follow-up visit to GPs who participated in the intervention.[13]

Research Question 2

Does the decreased utilization rate of COX-2 inhibitors for patients of GPs who have taken the AD intervention remain significant for a period of one-year post intervention? (Is the intervention effect sustainable?)

Research Question 2 hypothesis

The null hypothesis is that the OA AD intervention will not have a sustained effect on the decreased utilization rate of COX-2 inhibitors.

The alternative hypothesis is that the OA AD intervention will have a sustained effect on decreasing the utilization rate of COX-2 inhibitors.

The third research question examined whether patients of GPs in the intervention group experienced a change in the rate of medical services utilization due to a change in GI adverse events associated with traditional NSAID therapy (if there was a significant finding to research hypothesis 1). The hypothesis is divided into two categories: those that are related to pharmacotherapy and those that involve other medical services.

Research Question 3

Do patients of GPs who have undertaken the OA AD program have medical utilization rates associated with their OA that are significantly different from patients of GPs who have not participated in the intervention?

Research Question 3 hypotheses

The null hypothesis is that the OA AD intervention will have no effect on the utilization rate of (1) PPIs, (2) H2As, (3) misoprostol (4) GP office visits, (5) specialist office visits, and (6) death rates.

The alternative hypothesis is that the OA AD intervention will have the effect of changing the utilization rate of (1) PPIs, (2) H2As, (3) misoprostol (4) GP office visits, (5) specialist office visits, and (6) death rate.

The fourth research question examined whether one PS adjustment method was more successful adjusting for bias between groups based on measured bias reduction for

covariates that were not balanced after group assignment and the resulting sample size after PS methods were applied.

Research Question 4

Is there a superior PS method for the reduction of selection bias between the intervention and control groups?

Research Question 4 hypotheses

The null hypothesis is that there will be no difference in the three PS method's (quintile stratification, regression on the PS, and "greedy matching") ability to adjust for bias on unbalanced covariates.

The alternative hypothesis is that one PS method will adjust for bias on unbalanced covariates to a greater extent than the other two.

Significance of Research

This research is of significance to several groups within the healthcare system. The three groups that benefit directly from the research are patients, physicians, and health policy decision-makers. The results also add to the academic research in the area of effective behavioral change methodology and it adds to the methodology and understanding surrounding the use of PSs.

The largest impact of this research is in the area of health policy decision making. The decision to proceed with one course of action is often at the expense of others. This study will inform decision makers regarding the effectiveness of the OA AD intervention and allow them to make a more informed decision to continue with the AD detailing program to educate physicians on other health related topics or disease states.

This research adds to the validity of the research that has been accumulated in the area of AD. This is significant as it was concluded by Davis et al. in a systematic

literature review of AD that while AD is effective it is seldom used by providers of continuing medical education.[14] The uniqueness of this research lies in its analysis of a population based continuing AD program and not one that has been developed for the purposes of a single study.

This research advances PS methodology. It compares three PS methods in a real world and population based intervention. The results should contribute to the choice of PS methods employed by future researchers. The study also analyzes each of the propensity score method's ability to balance the control and intervention groups on unmeasured administrative variables. The ability of the propensity score methodology to balance groups on measured variables has been widely reported; however the ability of the methodology to balance unmeasured variables is assumed [15, 16] and studies attempting to measure the ability of the PS method to balance physician groups on a number of unmeasured administrative variables were not found in the literature.

CHAPTER 2
LITERATURE REVIEW

Literature reviews were conducted on two areas of interest: articles dealing with studies relating to AD interventions which have not shown statistical significance and articles relating to the use of PS methods. The AD studies which reported no statistically significant effects of AD interventions are of interest because they possibly give examples of shortcomings of methodology that may be of use in this study. The PS articles that are of interest to our study are those which involved studies that identified some unit, other than the patient, as the unit of analysis in the PS development and articles that dealt with PS methods.

The positive effect of AD on prescribing behavior has been summarized in a number of review articles on AD or educational outreach.[14, 17, 18] This body of evidence shows that AD moderately improves physician behavior and patient outcomes. Three review articles are summarized.

Review Articles Addressing Effects of Academic Detailing

Davis et al.[14] reviewed 99 studies which met their inclusion criteria from a total of more than 6000 articles. The 99 studies included 160 separate continuing education interventions, including academic detailing. Sixty-two percent of the interventions showed improvement in at least one major outcome with effect sizes ranging from small to moderate (quantified effect sizes not provided). There were fourteen AD interventions in the category of prescribing and 75% of these showed positive effects. AD was reported as an effective change agent for prescribing. The authors concluded that AD is

an effective strategy for continuing medical education (CME) however, it is not widely used by CME providers.

Thomson O'Brien et al.[18] conducted a systematic review of the effect of educational outreach on professional practice and health care outcomes. Eighteen studies were included in the review with thirteen of the studies targeting prescribing practices. Nine of the thirteen studies employed multifaceted interventions (educational outreach combined with reminders, audit and feedback, marketing, or patient-mediated interventions). Seven of the nine studies using multifaceted interventions showed statistically significant effects with relative effects ranging from 1 to 45% improvement (table 2-1). The authors noted that potential bias exists in thirteen of the eighteen studies due to lack of randomization and six of the studies contained potentially important baseline differences and adjustment for these differences was not carried out in the statistical analysis. It was also noted that only one of the eighteen studies considered patient outcomes. The authors concluded that the effects of educational outreach are small to moderate but potentially of practical importance.

Grimshaw et al.[17] conducted a systematic review of the effectiveness and costs of different guideline development, dissemination, and implementation strategies. 235 studies representing 309 comparisons were included in the review. The sections of the review that are germane to our study are the multifaceted comparisons involving academic outreach with continuous measures for process or outcome variables. Ten comparisons were reviewed which contained measures on continuous variables. Six of the comparisons involved process measures (five cluster randomized control trials and one controlled before and after trial) and all reported improvements in performance with

a median effect of 15.0% (range 1.7% to 24.0%) relative improvement. None of the studies included enough information to calculate standardized mean difference, and two studies were not statistically significant. Four of the comparisons involved outcome measures (three cluster randomized control trials and one controlled before and after trial). The median effect of the cluster randomized control trials was 0% (range -1.4 to 2.7%) and the standardized mean difference was calculated as 0 for one trail. The controlled before and after trial reported a relative improvement of 13.9% with a standardized mean difference of 2.38. The authors summarized the multifaceted interventions, including academic outreach, to be at best moderately effective (table 2-1).

Table 2-1. Summary of Included Studies for Thomson O'Brien and Grimshaw.

Author (year)	Reviewed by	Interventions (plus AD)	Relative Effect (%)
McConnell (1982)	Thomson O'Brien	Audit and Feedback (AF), Educational Material (EMat)	45.8
Stergachis (1987)*	Thomson O'Brien	AF, Patient Mediated (PM), Conferences	35.7
Meador (1997)	Grimshaw	EMat, Educational Meeting (EMeet)	24.0
Ross-Degnan (1996)	Thomson O'Brien	EMat, Social Marketing (SM), PM	21.0
Peterson (1996)	Grimshaw	EMat	20.0
Avorn (1983)	Thomson O'Brien	EMat, SM	15.2
Avorn (1992)	Thomson O'Brien, Grimshaw	EMat, SM, Conferences	15.0
Ray (1993)	Grimshaw	EMat, EMeet	13.9
de Burgh (1995)*	Thomson O'Brien	EMat, PM	13.0
Diwan (1995)	Grimshaw	EMat	11.3
Steele (1989)	Thomson O'Brien	Reminders	11.2
Santoso (1996)	Thomson O'Brien	EMat, SM	8.7
Schmidt (1998)	Grimshaw	Organizational Change	5.5
Elliott (1997)	Grimshaw	EMat, Opinion Leaders	2.7
Feder (1995)	Grimshaw	AF	0.0
Moore (1997)	Grimshaw	EMat, Reminders, PM	-1.4

* non-significant study results

Academic Detailing Studies Reporting No Statistically Significant Effect

Five articles were reviewed in which the authors reported non-significant results for AD interventions with pharmacotherapeutic outcomes. The review of results that were not positive is important because it will possibly indicate to investigators methodological similarities that may have been employed in previous unsuccessful studies. If identified the methodological shortcomings could be avoided.

Lin et al.[19] studied the effects of physician training on the management of depression. The study was a before and after design with an equivalent control group. The physician sample was made up of 109 primary care physician volunteers and they were associated with fifteen primary clinics. Randomization of groups was at the clinic level resulting in 56 physicians in the intervention group and 53 physicians in the control group. The intervention was outlined including the four key messages and the use of opinion leaders in intervention delivery.[20] Case managers were used for follow-up visits with the physicians. The intervention involved other components such as small group discussions, role-play and psychiatric consults. The authors reported that the physicians in the intervention arm of the trial did not differ significantly from the control group in adequacy of pharmacotherapy ($p=0.53$). While insignificant, the results showed a decrease of 7.5% in the percent of patients in the intervention group who received adequate pharmacotherapy with no change in the control group. The decrease in the intervention group is opposite to the desired outcome of the intervention. The study also failed to show significant differences in the number of antidepressant prescriptions per 100 patients ($p=0.10$). The percent of patients receiving new prescriptions in the intervention group decreased by 10.4% and increased in the control group by 4.8%. These results are opposite to the desired outcome of the intervention. The authors

reported that the study's main failure was its lack of power to detect a significant change between groups. The sample size used was sufficient to detect a 40% to 50% difference in adequate pharmacotherapy and a 15% to 30% difference in new antidepressant prescriptions. The fact that the effect of the intervention was the opposite of the hypothesis was not explained by the authors.

Brown et al.[21] studied the effect of AD and continuous quality improvement (CQI) interventions on the treatment of patients for depression. The study was a randomized controlled trial. The primary care clinician groups were randomized by first matching clinicians according to specialty (internal medicine or family practice), sex, training (physician or allied health clinician), and number of patients in a high-risk depressive cohort. The resulting sample size was 160 with 79 in the intervention arm and 81 in the control arm. The AD intervention involved focus groups for the collection of baseline knowledge of primary care providers (physicians, physicians' assistants, and nurse practitioners) in preparing the intervention. The intervention was based on guidelines from the Agency for Health Care Policy and Research and used the same material as the Goldberg study.[22] Three main messages were summarized on letter sized illustrated handouts. Four visits were used to deliver the message and the detailers were pharmacists from the clinicians' own medical office. The study showed mixed results. It was successful in increasing the percent of patients receiving antidepressant treatment (7.5% increase, $\underline{p}$=0.046 in depressed arm and 0.7% increase, $\underline{p}$=0.025 in the non-depressed population) however, it was not successful in increasing the total days of antidepressant therapy (16.7 days effect, $\underline{p}$=0.189 in the depressed arm and 1.3 days effect, $\underline{p}$=0.606 in the non-depressed population). The study did not exhibit significant

differences in non-pharmacotherapeutic outcomes (improvement of symptoms and measures of functional status). The authors report that the mixed findings could be due to the complexity of the implementation of a clinical guideline and the evidence base for the guidelines may not be generalizable to the study population. They propose that AD may be appropriate for behavioral change but is not sufficient for the implementation of clinical guidelines. This conclusion is important for our study since the primary outcome is change in prescribing behavior.

Goldberg et al.[22] studied the effect of AD and CQI interventions on compliance with guidelines for hypertension and depression. The study was a randomized before and after design with two experimental groups (AD only and AD combined with CQI) and an equivalent control group. The physicians were part of fifteen clinics and group randomization was carried out at the clinic level. The resulting sample size was 78 with 18, 37 and 23 physicians in the AD only, AD combined with CQI and usual care groups respectively. The AD intervention was based on national guidelines for hypertension and depression from the Agency for Health Care Policy and Research. Five recommendations were developed including two which specifically addressed pharmacotherapy. The AD intervention was delivered by opinion leader physicians and follow-up visits were conducted by staff pharmacists. The intervention was supported by handouts and pocket cards for quick reference. The study found significant effect in only one of the pharmacotherapeutic outcomes which was a decrease in the prescribing of 1st generation antidepressants to previously diagnosed depressed patients (relative effect -4.7%, $\underline{p}$=0.04). The other outcomes prescribing of hypertension medications, antidepressants to previously undiagnosed patients, 2nd generation antidepressants to

previously diagnosed depressed patients, and SSRIs to previously diagnosed depressed patients exhibited insignificant change with relative effect sizes and p-values of 1.3%, p=0.06; 2.4%, p=0.68; -2.1%, p=0.43; and 3.3%, p=0.11 respectively. One possible explanation for the failure of the study to show significant effect for all but one of the pharmacotherapeutic outcomes can be attributed to the presentation of too much information. A successful AD intervention should include only a limited number of messages regarding a disease state.[13] The presentation of an intervention covering two distinctly different disease states clearly violates this principle.

Zwar et al.[23] studied the effect of AD on prescribing rates of benzodiazepines for all indications. The study was a before and after design with an equivalent control group. There were 157 physicians who participated in the study. They were randomized into the benzodiazepine AD group (n=79) and the control group who received AD on another topic (n=78). The AD intervention was based on guidelines developed by the Royal Australian College of General Practitioners and it was delivered by physicians trained in AD techniques. The intervention was not accompanied by any other methods (i.e. handouts, etc.). The study found significant effect in overall prescribing of benzodiazepines (-26.7%, p=0.042) however, there was no significant between group relative effect (-1.2%, p=0.99). The authors attributed the lack of significant results to the effects of a pre-intervention practice survey that was given to all physicians in the study and a lack of power to detect a difference between groups due to the decision to aggregate data into eight subgroups thereby reducing the sample size dramatically.

Tomson et al.[24] studied the effects of AD on physicians' practice in the management of asthma and on patient knowledge. The study was not randomized and

the sample consisted of 63 GPs in two regions, one region was assigned as the treatment group (n=44) and the other region was assigned as the control group (n=19). The intervention was developed using existing physician knowledge as the baseline and the input of respirologists. It was delivered by a clinical pharmacologist and a pharmacist and contained three main messages. The face-to-face visits with physicians were augmented with written materials. The study found that there was not a significant difference between the treatment and control groups in prescribing ratios of beta-agonists and inhaled corticosteroids (no p-value reported). One explanation for the insignificant results could be attributed, at least in part, to insufficient power (due to the small sample size) to detect a meaningful change. The authors identified a possible selection bias in the physicians volunteering for the intervention as they may have been largely physician interested in asthma therapy to begin with.

It is important to note that a review of the negative findings of studies in the literature cannot be considered to be complete since many studies, and their fatal flaws, are not published if they are not considered to be methodologically sound or clinically important (publication bias). However, from the review of literature that did not report significant results there are four areas of inadequacy that the studies appear to have in common;

- the authors reported that there were insufficient sample sizes to yield enough power to show a meaningful change in the studies conducted by Lin[19], Gorins[25], Zwar[23] and Tomson[24], however in all but one of the studies[19] the effect of the intervention was consistent with the study hypothesis. It is important to note that lack of power is only one explanation for the lack of study significance,
- there were intervention development problems in that the interventions were too complex[21, 22],
- the interventions may have been compromised through the use of less than credible academic detailers[19, 25], and

- the use of pre-tests or pre-intervention surveys decreased the intervention effect due to a pre-sensitization of the subjects to the intervention.[23, 25]

The results from the above studies are applicable to our study for the following reasons. The lack of power reported by a number of the studies is only one explanation. Other explanations could include a large variation in measurement on the dependent variable or a lack of control for the variables that are associated with the outcome variable. For example, in our study we could have a large sample but if the number of elderly patients in the GP's panel is not controlled for then the variation could be inflated and a non-significant result could occur. In our non-randomized study design it is important to adjust for variables which are associated with the outcome but it must be acknowledged that there will be variables which are important confounders and are not measured so residual confounding (bias) will exist. There may be a need to adjust for patient variables as well. For example, if the GP's patient panel is markedly ill then this will confound the results. A measure of patient wellness would help to address this problem. The lack of a follow-up visit to GPs in our study may play an important role in the outcome.

Propensity Scores

A literature search was conducted using PubMed for all years up to and including July 20, 2005. The search terms used were "propensity score" and "propensity scores". The search yielded 341 articles. The abstracts for all 341 articles were reviewed and the distribution of articles by article objective and year is illustrated in figure 2-1.

The distribution shows an initial surge of articles dealing with PS methods in the late 1990's with articles containing objectives other than medical (e.g., economic) and only a few articles with stated medical objectives. Since 2000 there has been a surge in

published articles using PSs particularly in the field of cardiology. The increase in use of PSs has been mirrored by an increase in published articles dealing with PS methods.

There were four articles which used a unit of analysis for PS other than the patient. Two of the articles used human couples as the unit of analysis one article developed PS on hospitals and one article used communities as the unit of analysis for the PS. There were no articles found which used the physician as the unit of analysis in the PS analysis.

There were 50 articles that described PS methods and 24 of these were selected for further review. Criteria for selection included PS studies using GEE for outcome models, studies comparing small experimental groups, studies describing PS and sample sizes or studies which described PS methods in detail. The information gained from these articles plus reference material from previous course work, library searches, and colleagues formed the basis for the PS method as it has been applied in this study.

In our study the PS represents the probability of a physician volunteering for the OA AD intervention given a number of personal and practice characteristics. For studies using quasi-experimental designs it is important to include methods to compensate for the lack of randomization to experimental groups. In our study we have made multiple measures of outcome variables both before and after the intervention and we have included a control group for comparison. The control group is not equivalent to the intervention group so adjustment on PSs was used to reduce the effect of the between group bias.

Three methods for applying PS in observational studies are predominant in the literature.[26] The three methods are; sub classification on the PS[11, 12, 27], regression on the PS[12] and matching on the propensity score using Mahalanobis metric matching[12, 27] or

"greedy matching" techniques.[28] All three of the methods; stratification, regression on the PS, and matching have been applied successfully in observational studies and therefore all three will be considered for application in this research.

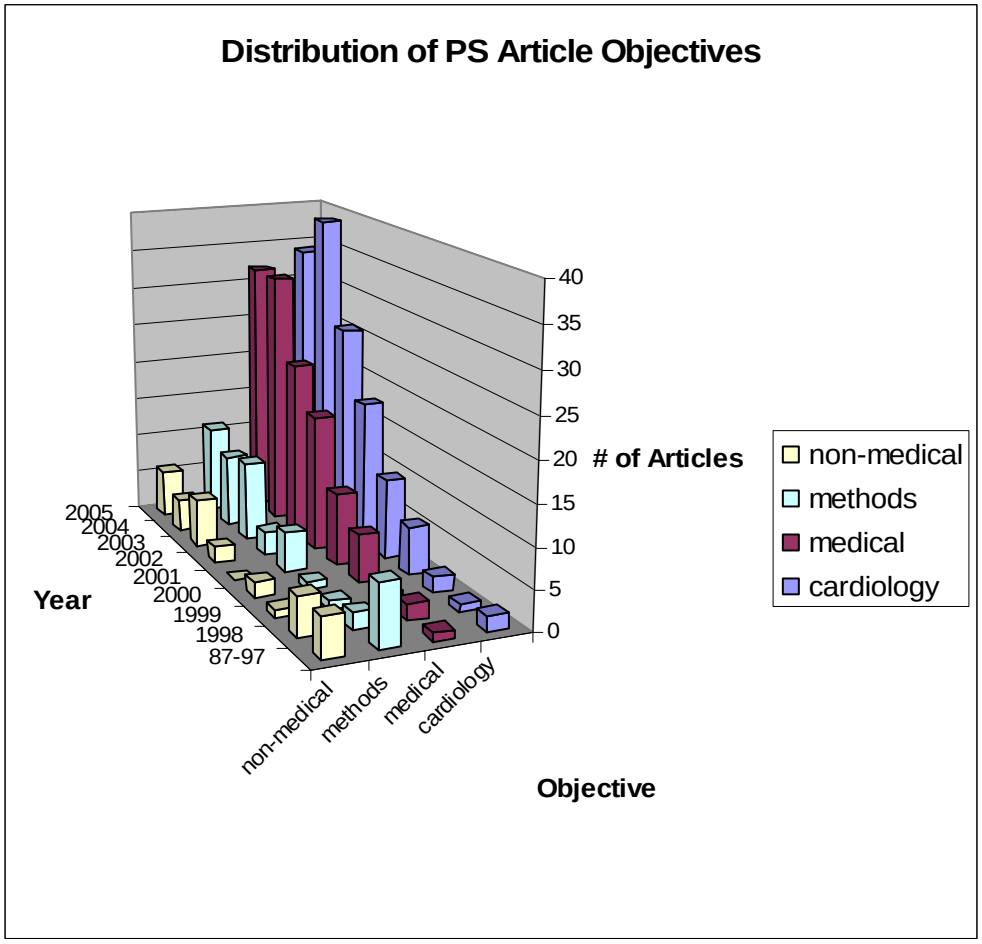

Figure 2-1. Distribution of Propensity Score Article Objectives: 1987 to July 20, 2005

An overarching limitation of all three PS methods is that the PS can only adjust for bias in observed covariates[27] and the extent to which the bias is abated in unobserved covariates depends on the correlation of the unobserved covariate with one that is observed.[11] Shadish stated that if the PS method was successful in abating bias in the measured covariates then the assumption can be made that the methodology would be successful in decreasing the bias in unmeasured covariates as well.[16]

A recent study has tested the ability of PS based on covariates extracted from administrative data to reduce the bias in unmeasured clinical variables. In this study the clinical data was extracted from patients' charts after PS methods were applied. The experimental groups set by the PS method were tested for significant difference on the clinical variables and it was found that the clinical variables were not balanced between the groups.[29]

Other studies have explored the number of events per variable that are needed for logistic regression analysis to outperform the PS method. Cepeda et al. reported that in their simulation model if there are six or fewer events per independent variable (covariates in the PS model) then the PS estimates are less biased then the regression estimates.[30] It is important to note that even if the number of subjects exceeds six the use of PS methods is warranted since it is a variable which predicts the exposure of interest without including the outcome[11] and that the use of PS methods is intended to complement model-based procedures not replace them.[31]

There are two measures of PS model fit that are reported in studies. The c-statistic is the area under the receiver operating characteristics (ROC) curve and is a measure of the discriminative ability of the PS model.[32, 33] The range of the statistic is from 0.5 to 1.0. If a model has a c-statistic of 0.80 this can be interpreted as the model accurately assigning random pairs of subjects to their experimental groups based on PS alone 80% of the time. The c-statistic is intended to be an indicator in the model building process but it is not a measure of the PS model's ability to adjust for bias[15] and it has not been found to be associated with the ability of a PS model to reduce residual confounding.[32] The goodness of fit is another statistic that is commonly used in regression analysis. Like

the c-statistic these tests were not found to be useful in predicting the ability of the PS model to reduce residual confounding.[32] As a result, these measures were not used in our study to decide which PS method to use for the outcomes analyses. The c-statistic was however, used to explain the effects on the model's discriminatory ability when variables were intentionally removed from the PS Regression model.

CHAPTER 3
METHODS

This study is a retrospective cohort, before and after longitudinal design with a non-equivalent control group using the Nova Scotia Medical Services Insurance and the Canadian Institutes of Health Information datasets for analysis. The non-equivalent control group design requires the use of procedures to abate selection bias in the treatment group.[12]

The methodology for the study can be broken down into four distinct sections, which are as follows;

- the extraction and validation of data from the administrative databases,
- the establishment of balanced control and experimental groups using three distinct PS methods,
- the primary outcome analysis of the intervention effect on the utilization rate of COX-2 inhibitors and,
- the secondary outcome analysis of the intervention effects on the utilization of PPIs, misoprostol, and H2As.

Step One: Extraction and Validation of Data

Sources of Data

All of the data used in this study was collected in pre-existing administrative databases. There were no occurrences of missing data since the variables included in the analysis were extracted from long standing registrar data which is complete for all fields listed in the registry[1] (GP demographics), complete census information[2] (geographic data) or the data was reported in terms of rates with the GP inclusion criteria ensuring

that each GP panel contained at least twenty patients so the rates for the outcomes measures were always defined (i.e. rate denominators were not equal to zero). Administrative data must be used with caution as it is not 100% reliable. Chapter five outlines the limitations of the administration used in our study.

GP demographic data for all GPs in the province was obtained from the Nova Scotia College of Physicians and Surgeons Physician Registry (2002).[1] The Dalhousie CME Division provided data which contained demographic information of the GPs who were detailed and the dates when the detailing visits were carried out. These two sources of data were merged and the resulting file was submitted for encryption using the same encryption methods as the provincial administrative data. The resulting encrypted GP demographic profiles were augmented with data from the Nova Scotia Medical Services Insurance (MSI) physician registry (2002) to include dates indicating when the GPs opted in and opted out of the provincial pharmacare billing scheme. GP practice information such as population of the community and average income of the county in which the practice is located was added to the demographic profile of each GP.[2]

Patient data was extracted from the Nova Scotia Pharmacare Seniors Dataset (2002-2004) and the hospital discharge data found in the Canadian Institute of Health Information (CIHI) hospital discharge dataset (2002-2003). Patient level GP visit data was used to determine to which GP's patient panel a patient belonged (see patient inclusion criteria). Once the patients were assigned to GP panels the patient prescription claims data and hospital length of stay data were aggregated at the GP level. Drug utilization variables were created at the GP level with the unit of measure equal to DDDs per elderly patient per 90 day study period. Change in utilization rate variables were

created for each GP by subtracting each period utilization rate (period = 1 to 6) from the baseline (period two) utilization rate. Period two was chosen as the baseline utilization rate since it was the pre-intervention measure most proximal to the GP index date.

Descriptive statistics for the GP demographic variables were calculated to confirm that the variables did not contain any missing data and to confirm that the variables fell with acceptable ranges (i.e. no GPs 200 years old, not all male GPs). The descriptive statistics are reported in tables 4-1 and 4-2.

Prescription claims, GP visits and vital statistics were checked to ensure that there were not instances of missing data. The prescription claims and GP visit data were complete on all fields necessary for our study. Only hospital admissions and deaths due to GI events were included in the hospital length of stay and death measures. A detailed description of the inclusion criteria for data is contained in chapter four. The underlying and primary causes of death were used to determine death rates and cause of death and data was reported for all included patients who died over the study period. The first four diagnoses codes for hospital admission were used to determine if GI complications were associated with admission. In all cases there was at least a primary diagnoses on admission. While the data for our study was complete it was administrative data and there are shortcomings associated with it. The limitations of administrative data are described in chapter five.

Data from several administrative databases was linked to create the datafile necessary for the PS analysis and for the outcomes analysis. The data linkage was carried out using the encrypted physician identifiers and the encrypted patient identifiers. The

encryption of the patient and physician identifiers was carried out according to standards set by the Canadian Institutes of Health Information (CIHI).[34]

GP Inclusion Criteria

The academic detailing intervention was targeted at GPs and, therefore, the experimental unit is the GP and the patient data for the GP's practice is the unit of measure. Each GP's practice is measured as an aggregate of the individual patient's data from his or her practice. The aggregation of patient data is described in greater detail later in this chapter. The date on which the GP received the OA AD intervention was defined as the index date. For GPs in the control group the index-date was randomly assigned from the time period over which the AD intervention took place.

There are four criteria that a GP had to meet to be included in the study. They are as follows;

- The GP had to be registered as a GP with the Nova Scotia College of Physicians and Surgeons for the entire study period.
- The GP had to be included on the billing registry with the Nova Scotia Medical Services Insurance (MSI) (the provincial government payment agency for seniors' medical and pharmacy claims) for the entire length of the study. This registry is the source of the medical and pharmacy claims data that will be used for the outcomes analysis.
- The GP had to have an elderly patient panel equal to or greater than twenty patients. The rational for the cut score of twenty was based on the premise that a 5% decrease in COX-2 utilization (i.e. COX-2 utilization rate change from 6.0% to 5.7%) will equate to annual savings to the elderly population of approximately $100,000. Therefore, if the GP had an elderly patient panel equal to twenty he or she was required to change prescribing behavior for one patient over the study period to realize a 5% change.
- The GP had to have at least one prescription claim for a COX-2 inhibitor recorded in the pre-intervention period (6 months preceding the GP's index date).

Patient Inclusion Criteria

The patient is the unit of measure for this study. Patients had to meet two criteria for inclusion in the study. The criteria are:

- The patient had to be included on a GP's patient panel. For inclusion on a GPs panel the patient must have seen a specific GP for more that 50% of his or her total GP visits for the fiscal year ending March 31, 2002. For example, if a patient had a total of forty GP visits in the period from April 1, 2001 to March 31,2002 and twenty-four (60%) of the visits were billed by one GP the patient was included on the GPs patient panel. Once the patient was assigned to a particular GP they remained with that GP throughout the study.

- The patient had to be 66 years of age or older as of the GP's AD index date. This ensures that the patient was at least 65 years old and eligible for the MSI pharmacare coverage for the entire study period and it provides a period of time of at least six months for the patient to become accustomed to the new MSI pharmacare coverage.

Step Two: Adjustment for Confounding Using Three Distinct Propensity Score Methods

The definition of the PS is the conditional probability of treatment given the individual's covariates. In this case it would be the conditional probability of taking the OA AD intervention given the GP's personal and practice characteristics.

The PS is obtained by fitting the data using a logistic regression model.[5] Once the PSs were calculated for each GP three PS methods were applied to the PS data and the optimal method in term of bias reduction and resultant sample size was determined.

The three PS methods used in this study were; the stratification into quintiles, regression on the propensity score[12], and "greedy matching" or one-to-one matching for group assignment.[28]

The variables in the regression model describe the GPs' personal characteristics (age, sex, birthplace, etc.) and practice characteristics (size of patient panel, population of community in which the practice is located, etc.). All variables in the data that fit within

these two descriptive categories were included in the regression model. This approach is consistent with the literature which calls for the inclusion of all variables which have some relevance to the outcome variable.[16] A description of the included variables and the abbreviations used in our study are included in table 3-1.

Table 3-1. PS Model Variable Descriptions and Abbreviations

Variable Description	# Levels	Abbreviation
GP participation in the OA ADintervention	2 (Y/N)	OA
GP participation in previous influenza AD service	2 (Y/N)	flu AD
GP's sex	2 (M/F)	sex
GP's birthplace	3 (Nova Scotia, Canada, Other)	birth place
GP's location of initial licensure	5 (Nova Scotia, Canada East, Canada Center, Canada West, Other)	license
GP's COX-2 utilization rate at baseline (DDDs / patient)	continuous	BL rate
GP's age (years)	continuous	GP age
population size of community in which GP's practice is located	continuous	population
average income of county in which GP's practice is located ($cdn)	continuous	aver income
number of patients in the GP's practice	continuous	total # pt
percent of GP's patients diagnosed with OA (ICD-9 CM = 715)	continuous	% OA dx
percent of GP's patients > 65 years old	continuous	% elderly
average hospital length of stay for elderly patients in the GP's practice (days/patient)	continuous	los rate

A logistic regression model was used to accommodate the dichotomous nature of the outcome variable, OA. The same regression model was applied using PROC REG (SAS 8.2)[35] for all three methods to determine GP PSs. Models described in this study

have categorical variables listed as single entities which are consistent with the SAS coding techniques. The model analysis creates (t-1) dummy variables (where t = the number of levels) for each categorical variable. The PS regression model is shown in figure 3-1.

$$Y = \alpha + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6 + \beta_7X_7 + \beta_8X_8 + \beta_9X_9 + \beta_{10}X_{10} + \beta_{11}X_{11} + \beta_{12}X_{12}$$

Where;
Y - GP participation in the intervention (0 = no, 1 = yes),
X_1 - GP participation in previous influenza AD service (0 = no, 1 = yes),
X_2 - GP's sex (Male, Female)[1],
X_3 - GP's birthplace (Nova Scotia, Canada, Other)[1],
X_4 - GP's location of initial licensure (Nova Scotia, Canada East, Canada Center, Canada West, Other)[1],
X_5 - the GP's COX-2 utilization rate at baseline (DDDs / patient).
X_6 - GP's age (years)[1],
X_7 - population size of community in which GP's practice is located[2],
X_8 - average income of county in which GP's practice is located ($cdn)[2],
X_9 - number of patients in the GP's practice,
X_{10} - percent of GP's patients diagnosed with OA (ICD-9 CM = 715),
X_{11} - percent of GP's patients > 65 years old,
X_{12} - average hospital length of stay for elderly patients in the GP's practice (days / patient).

Figure 3-1. Propensity Score Logistic Regression Model

Variables were kept in the model regardless of their significance. Variables that are not statistically significant still contribute to the model and the population based nature of the data ensures a large enough sample size to support the model with twelve predictor variables. The final model predicts the probability that each GP would receive the intervention based on his or her individual variables. This probability is the GP's PS.

Once the PSs were calculated they were applied according to the three methods stated earlier.

Quintile Propensity Score Method

For the quintile method; the GPs in the treatment and intervention groups were stratified, based on their participation in the OA AD intervention, and then ordered according to the GP's PS. The treatment and control groups were stratified into five levels, or quintiles. Each quintile contains 20% of the GPs (table 4-4).

Regression on the Propensity Score Method

For the regression on the PS method; the PS was used in the outcome model.

"Greedy Matching" Method

For the "greedy matching" method; the GPs in the treatment and intervention groups were stratified, based on their participation in the OA AD intervention, and then ordered according to the GP's PS. A matching procedure was applied[28] that involved matching the groups on PS beginning with matches accurate to five decimal places and concluding with matches to one decimal place. The number of included GP only allowed for a one-to-one match between groups. Once matched the GP was removed from the sample pool. Those GPs that were not matched were deleted. The "greedy matching" method resulted in group sizes of 104 each (N = 208 total).

Propensity Score Method Selection

The regression on the PS method was selected for use in the outcomes analysis. The regression on the PS method was selected based on the following criteria; the adjustment for selection bias on the covariates measured before and after the PS procedure is carried out, and the resultant sample size.

The adjustment for selection bias after application of PSs was determined for each PS method using the following methods.

For continuous variables the percent decrease in bias was calculated using the formula:[11] 100 x [1 - (bias post) / (bias pre)], where bias post was the difference between PS adjusted group means and bias pre was the difference between unadjusted group means (group means before PS analysis). Variable means before PS analysis are reported in table 4-3 as the unadjusted means of the groups. Variable means after PS adjustment for the regression on PS and quintile methods were the least square means reported using PROC GENMOD (SAS 8.2)[35] after adjustment for propensity score (or quintile depending on the method). For the "greedy matching" method unadjusted means were used for both the pre and post means calculations. Results are reported in tables 4-5, 4-7, and 4-8.

For categorical variables the percent decrease in bias was calculated using the following formula:[12] 100 x [1 – |(1- OR post) / (1- OR pre)|] where OR post is the odds ratio of the groups (adjusted for PS) and OR pre is the odds ratio of the groups before PS adjustment. For both the pre and post odds ratio measures PROC GENMOD (SAS 8.2)[35] was used. The odds ratios were calculated using the same procedure for all three PS methods. Results are reported in tables 4-5, 4-7, and 4-8.

A further test of the effect of the different PS methods involved the purposeful removal of independent variables from the regression model and the subsequent test for PS adjustment on the "unmeasured" variable. The logistic regression model was run twelve times. Each time one of the independent variables was removed from the model and the percent bias reduction on the now "unmeasured" variable was calculated for each of the three PS methods. The same equations for continuous and categorical variables

were used to calculate percent bias reduction on the variable that had been removed. The results are reported in tables 4-9 to 4-11.

The measurement of adjustment for bias on "unmeasured" variables was not considered in the selection of the PS method. It has been included in this study as a means of contributing to the PS methodology. Work has been done on the PS model's ability to adjust for bias on unmeasured clinical variables[29] and the PS model's ability to adjust for bias on unmeasured variables in a large computer generated dataset.[32] Our study is unique, however, since it examines the PS model's ability to adjust for bias on demographic variables contained in a relatively small, real world dataset.

There were five PS model covariates that showed significant between group differences after the initial OA group assignment. The variables were percent of patients diagnosed with OA (% OA dx), the average income of the county in which the GP's practice is located (aver income) the average hospital length of stay per patient (los rate), the population size of the community in which the GP's practice was located and participation in a previous influenza AD intervention (flu AD). The PS adjustment on the flu AD variable was not successful for any of the three PS methods so it was included in the outcomes models as a covariate. The other four variables were of interest in the analysis of effect of PS method's ability to adjust for bias on unmeasured administrative variables. The correlations between the variable and the PS were calculated and graphed against the percent reduction in bias for each PS method. Correlations between the variables and the included PS model covariates were calculated and tabulated. The relationship between the reduction in bias in unmeasured variables and each PS methods was studied. The results are contained in chapter four.

Step Three: Primary Outcome Analysis; Intervention Effect on COX-2 Utilization Rates

Once the method of PS analysis was selected and the GP intervention and control groups had been determined, the analysis of the primary outcome effect was carried out as described below.

To enable the analysis of the changes in COX-2 utilization rates over time the COX-2 utilization rates were determined for each GP in the study for six consecutive ninety-day time periods. Two time periods were pre-intervention and four time periods were post-intervention.(Figure 3-2) The index-date is reported as the date that the GP received the AD intervention and the index-dates for the control group were assigned by randomly selecting dates from the range of time that the AD intervention spanned.

The COX-2 utilization change rate will be calculated by subtracting the GP's baseline utilization rate (period 2 utilization rate) from the utilization rates in each study period.

Intervention Group	O	O	X	O	O	O	O
Control Group	O	O		O	O	O	O
Time from intervention (days)	-180 to -91	-90 to -1	Index date	1 to 90	91 to 180	181 to 270	271 to 360

Figure 3-2. Experimental Design Timeline

Before the utilization rates could be calculated the inclusion and exclusion criteria for claims in a given time period were defined. An example of the operationalization of the decision rules for the inclusion or exclusion of claims within a time period is presented using a fictitious ninety day time period (January 1st until March 31st) and describing how different scenarios were adjudicated. If a prescription claim is submitted

for a two-month supply on January 2nd it is clear that the period of time for the entire claim falls within the given time period and the claim is included. If a prescription claim is submitted for the same two-month supply on March 28th it is clear that the entire claim period does not fall within the period ending March 31st. In this case the claim would still be counted, in its entirety, in the claim period that it was submitted. The reason for inclusion of the claim in the initial time period is that it was in this time period that the GP's prescribing behavior took place and the intention was to have the patient take the medication as prescribed.

Refills were considered to be an extension of the original claim until such a time as the refill claim was submitted more than thirty days after the intended fill date for the refill. If the refill was more than thirty days late the rest of the claim was not counted in any time period.

The COX-2 utilization rates for each GP was determined through the use of the World Health Organization's (WHO) Anatomic and Therapeutic Classification System/ Defined Daily Dose (ATC/DDD) methodology[36] and was reported for each GP as the average number of DDDs per included patient per ninety-day intervention time period. The reporting of DDDs is often given as per thousand patients, however, since most GPs in the study will not have one thousand patients that meet the criteria this could be misleading.

DDDs are drug consumption data that are independent of price and formulation. Once set, the WHO is reluctant to change DDD measures and as such the DDD is stable over time. This makes the DDD measure more reliable for drug consumption studies but it is not appropriate for clinical analysis. The DDD, therefore, "enables the researcher to

assess trends in drug consumption and to perform comparisons between population groups."[36]

An analysis of the intervention effect on the primary outcome, change in COX-2 utilization rates, was carried out. The primary outcome model initially included the dependent variable (change in COX-2 rates for the four post-intervention periods), the independent variables indicating the between group effect (OA) and longitudinal effects (period), the PS variable, the variable flu AD (as indicated from the PS analysis), as well as baseline COX-2 rate (BL rate), and number of elderly patients in the GP's panel (# elderly pt). The model is depicted in figure 3-3. Each of the variables was retained in the model regardless of its significance. The covariates were all included as adjustments for confounding which if not controlled would be questioned in peer review. The included variables for the primary outcome model with their associated coefficients and significance levels are reported in table 4-13 in the results section.

$$\mathbf{Y = \beta_0 + \beta_1(X_1) + \beta_2(X_2) + \beta_3(X_3) + \beta_4(X_4) + \beta_5(X_5) + \beta_6(X_6)}$$

Where;
Y = change in COX-2 utilization rate (periods 3 to 6 (post-intervention)),
X_1 = GP participation in the intervention (0 = no, 1 = yes),
X_2 = experimental time period (period = 3,4,5,6),
X_3 = PS,
X_4 = GP participation in the influenza AD service (0 = no, 1 = yes),
X_5 = GP baseline COX-2 rate (DDD / patient / period = 2),
X_6 = number of GP's patients > 65 years old,

Figure 3-3. Primary Outcome Model for Between Group Effect

The model determined the statistical significance of the between group effect (between group effect) as well as the longitudinal effect (within subjects effect) of the intervention. The model was analyzed using PROC GENMOD (SAS 8.2)[35] and significance is reported at the alpha= 0.05 level.

The two ninety-day pre-intervention utilization rates for the experimental groups were analyzed to determine if there were any significant between group differences in the change in COX-2 utilization occurring before the study commenced. The pre-intervention analysis was carried out using the same model as the primary outcome model described in figure 3-3 however, only the first two time period measurements (period = 1,2) for each GP were entered into the model. This examined whether or not significant differences for change in utilization rates were present between the groups before the intervention was applied.

A longitudinal model was tested using the primary outcome model in figure 3-3 with a pre-intervention / post-intervention variable added which described whether the change in COX-2 utilization rate was pre- or post-intervention. The measurements for periods one and two were coded as prepost =1 and the measurements for periods three to six were coded as prepost = 2. The longitudinal effects model was run twice with only one of the intervention groups included each time. The prepost variable indicated whether a significant within group intervention effect occurred. The results are reported in tables 4-18 and 4-19.

Step Four: Secondary Outcome Analyses; The Utilization of Other Health Care Resources Associated with NSAID Induced GI Side Effects

The primary outcome model exhibited significant between group differences and therefore, all of the secondary outcome analyses were carried out using the same GP groups as in the primary outcome analysis. The models for the secondary outcomes were developed using the same variables as the primary outcome model. Each secondary outcome model had the change in COX-2 utilization rate substituted with the appropriate secondary outcome rate. The secondary outcomes that were analyzed are the intervention

effect on changes in rates from baseline for; PPI utilization, misoprostol utilization, H2A utilization, GP office visits, specialist office visits, and death rates due to GI complications. Rates for secondary outcomes (described individually with each outcome analysis) and the results for the secondary outcomes are described and reported in chapter four.

CHAPTER 4
RESULTS

Step One: Extraction and Validation of Data

PROC MEANS[35] was used to perform the calculations for the continuous variables. The variable mean, standard deviation, median, minimum and maximum are reported in table 4-1.

Table 4-1. Descriptive Statistics for Continuous Variables in the PS Model

Group	N	Variable	Mean	Std Dev	Median	Minimum	Maximum
AD= 0	265	% OA dx	0.0913	0.1071	0.0638	0.0000	0.7391
		GP age	47.30	9.78	47.00	27.00	79.00
		% elderly	0.1910	0.1160	0.1676	0.0253	0.9000
		total # pt	1054.18	438.55	1021.00	30.00	2575.00
		aver income	27688.68	4683.37	27500.00	22500.00	32500.00
		BL rate	3.6013	2.7048	3.0675	0.0769	16.8750
		los rate	0.0453132	0.1449	0.0000	0.0000	1.2647
		population	183342.23	162415.32	109330.00	991.00	359183.00
AD= 1	231	% OA dx	0.0719	0.0598	0.0608	0.0000	0.2601
		GP age	45.74	9.18	45.00	27.00	77.00
		% elderly	0.1772	0.0788	0.1681	0.0251	0.5564516
		total # pt	1037.69	418.99	1009.00	171.00	2481.00
		aver income	25833.33	4488.31	22500.00	22500.00	32500.00
		BL rate	3.9758	2.8707	3.4456	0.0487	14.2500
		los rate	0.0882	0.1767	0.0000	0.0000	1.2592
		population	122705.25	155567.12	22430.00	550.00	359183.00

PROC FREQ[35] was used to perform the calculations for the categorical variables. The proportion of each variable level is reported in table 4-2.

Step Two: Establishment of Balanced Control and Experimental Groups Using Three Propensity Score Methods

Pre-Propensity Score Analysis

Twelve variables were identified in the administrative data as describing personal and practice characteristics of GPs. GP age, sex, participation in a previous influenza AD

intervention, place of initial licensure, baseline COX-2 prescribing rate, birthplace and percent of patients diagnosed with OA describe personal characteristics. The percent of elderly patients, total number of patients, average income of county where the practice is located, population size of the community where the practice is located and average hospital length of stay for patients describe the GP's practice characteristics.

Table 4-2. Descriptive Statistics for Categorical Variables in the PS Model.

Variable	Level	Proportion	
		AD = 0 (n=265)	AD = 1 (n=231)
Sex	Female	0.3057	0.2987
	Male	0.6943	0.7013
Flu AD	Yes	0.1585	0.7273
	No	0.8415	0.2727
License	Canada	0.1208	0.1732
	Nova Scotia	0.6717	0.6623
	Other	0.2075	0.1645
Birth place	Nova Scotia	0.4830	0.4545
	Canada East	0.1245	0.1255
	Canada Centre	0.0679	0.0736
	Canada West	0.0264	0.0390
	Other	0.2981	0.3074

Table 4-3 contains the pre-PS variable values which include; means and standard deviations for continuous variables, F-statistics (square of t-test for continuous variables and F for coefficient estimate from PROC GENMOD[35] for categorical variables), p-values, coefficient estimates for the main effect of the intervention for the categorical variables, and the odds ratio for the main effect. These values were used in subsequent tables to calculate percent bias reduction for each PS technique.

The pre-PS analysis indicates that there are five variables that are not balanced. These variables are of the greatest concern since the goal of PS methods is to balance the groups on measured covariates.[12, 16] The five variables that show significant differences at the alpha = 0.05 level will be collectively referred to as the variables of concern (VOC)

and they are; the percent of elderly patients with a diagnosis of OA (% OA dx), the average income of the county in which the physician's practice is located (aver income), the average hospital length of stay rate for elderly patients per physician (los rate), the population of the community in which the physician's practice is located (population), and physician participation in a previous influenza AD service (flu AD).

Table 4-3. Pre-PS Univariate Analysis for Included Variables.

Pre-Propensity Score Values (t-test and proc genmod)								
	AD = 0 (n = 265)		AD = 1 (n = 231)					OR
Variable	mean	std dev	mean	std dev	F	p - value	B	(exp B)
% OA dx*	0.0913	0.1071	0.0719	0.0598	9.9191	0.0116		
GP age	47.3	9.8	45.7	9.2	9.2191	0.0678		
% elderly	0.1910	0.1160	0.1772	0.0788	8.9591	0.1182		
total # pt	1054	439	1038	419	7.8191	0.6700		
aver income*	27689	4683	25833	4488	11.8791	0.0001		
BL rate	3.60	2.70	3.98	2.87	5.8991	0.1356		
los rate*	0.0453	0.1449	0.0882	0.1767	4.4191	0.0031		
population*	183342	162415	122705	155567	11.6191	0.0001		
sex					0.0300	0.8663	-0.0330	0.9675
flu AD*					140.1500	0.0001	-2.6503	0.0706
license					3.3600	0.0669	0.3460	1.4134
birth place					0.1100	0.7415	-0.0553	0.9462

* variables that show significant differences at the alpha = 0.05 level

Quintile PS Method Analysis

The distribution of GPs within the quintiles is reported in table 4-4. Quintile one represents the GPs with the lowest PSs (lowest propensity to volunteer for the intervention) and quintile five represents the GPs with the highest PSs (highest propensity to volunteer for the intervention). The table is consistent with the expected PS distribution with fewer subjects in the high propensity quintile for the control group and fewer subjects in the low propensity quintile for the intervention group.

The results for the quintile method were generated using PROC GENMOD[35] and are reported in table 4-5. The main effect column represents the main effect of the AD

variable and the interaction effect column represents the effect of the AD by quintile interaction. The quintile method resulted in no statistically significant difference between groups on all five VOC while maintaining balance on the rest of the covariates.

Table 4-4. Physician Distribution by Quintile

Quintile #	Intervention	Control	# of GPs
1	86	13	99
2	77	22	99
3	65	35	100
4	24	75	99
5	13	86	99
TOTAL	265	231	496

Table 4-5. Quintile Method Regression Analysis Results.

	Quintile Method								
	lsmean		Main Effect		Interaction Effect				
Variable	AD = 0 (n = 265)	AD = 1 (n = 231)	F	p	F	p	B	OR (exp B)	% bias reduction
% OA dx*	0.0827	0.0783	0.20	0.6562	1.74	0.1403			77.32
GP age	47.2	47.1	0.01	0.9313	1.01	0.4029			93.75
% elderly	0.1918	0.1808	0.93	0.3351	2.04	0.0880			20.29
total # pt	1045	1031	0.09	0.7673	0.22	0.9256			12.50
aver income*	26757	26693	0.02	0.8900	1.34	0.2526			96.55
BL rate	3.58	3.68	0.10	0.7479	0.90	0.4616			72.89
los rate*	0.0491	0.0677	1.05	0.3058	0.76	0.5529			56.64
population*	149548	148905	0.00	0.9670	1.27	0.2797			98.94
sex			0.00	0.9818	0.00	0.9801	-0.0130	0.9871	60.21
flu AD*			0.35	0.5537	xx[+]	xx[+]	-0.3723	0.6891	66.55
license			3.26	0.0709	3.28	0.0700	-0.9702	0.3790	-50.22
birth place			0.01	0.9372	0.01	0.9295	-0.0381	0.9626	30.51
Average**									82.36

* variables that were not significant at the alpha = 0.05 level in the pre-PS model

** average % bias reduction for variables with significant differences in the pre-PS model (excluding flu AD)

[+] estimates not available (see table 4-6 for explanation)

The interaction effect (AD*quintile) was not significant for four of the five VOC however, the flu AD variable exhibited an almost complete separation of data points (table 4-6) and as such the interaction effect was not estimated.

The distribution of the flu AD variable on the PS was problematic for all three PS methods (figure 4-1). Therefore, the reported average percent bias reduction on the VOC

does not include the flu AD variable. The average percent bias reduction for the quintile method is 82%.

It is evident at this point that the flu AD variable will have to be included in the outcome models regardless of the PS method chosen.

Table 4-6. Distribution of Influenza AD Participants by Propensity Score Quintile

Flu AD Participation	**Quintile**					**Total**
	1	2	3	4	5	
No	99	99	88	0	0	286
Yes	0	0	12	99	99	210

Regression on the Propensity Score Method Analysis

The results for the regression on the PS method were generated using PROC GENMOD[35] and are reported in table 4-7. This method was successful in balancing three of the five VOC while maintaining balance on the rest of the covariates. The average percent bias reduction on the VOC (flu AD excluded) is 99%.

The variable, population, retained a significance level less than 0.05 and it also exhibited a significant interaction effect (population*AD) at the alpha = 0.05 level. The variable aver income showed a non-significant main effect with a p value > 0.05 however, the interaction effect (aver income*AD) is less than the 0.05 level. The separation of data points for the flu AD variable on the PS was again evident. Figure 4-1 shows the distribution of flu AD on PS (stratified at 0.05 intervals). This separation precluded the model from estimating main and interaction effects for flu AD.

"Greedy Matching" Method Analysis

The results for the "greedy matching" method were generated using PROC GENMOD[35] and are reported in table 4-8. This method was successful in balancing four

of the five VOC while maintaining balance on the remainder of the covariates. The average percent bias reduction on the VOC (flu AD excluded) is 75%.

Table 4-7. Regression on PS Method Analysis Results.

Regression on Propensity Score Method

	lsmean		Main Effect		Interaction Effect				
Variable	AD = 0 (n=265)	AD = 1 (n=231)	F	P	F	p	B	OR (exp B)	% bias reduction
% OA dx*	0.0770	0.0772	2.19	0.1394	3.07	0.0802			98.97
GP age	47.0	46.9	0.90	0.3441	1.30	0.2548			97.50
% elderly	0.1819	0.1819	0.45	0.5030	0.63	0.4291			100.00
total # pt	1036	1037	0.33	0.5637	0.47	0.4924			93.75
aver income*	26531	26530	2.83	0.0934	3.92	0.0482			99.95
BL rate	3.68	3.69	0.61	0.4338	0.89	0.3462			97.63
los rate*	0.0613	0.0620	0.32	0.5697	0.48	0.4872			98.37
population*	142500	142661	4.23	0.0402	5.90	0.0155			99.73
sex			0.22	0.6392	0.31	0.5746	0.2073	1.2304	-609.62
flu AD*			xx+	xx+	xx+	xx+	728.77		
license			2.48	0.1152	3.48	0.0622	-0.6744	0.5095	-18.66
birth place			0.03	0.8588	0.04	0.8377	-0.0674	0.9348	-21.15
Average**									99.25

* variables that showed significant differences at the alpha = 0.05 level in the pre-PS model

** average % bias reduction for variables that were not significant in the pre-PS model (excluding flu AD)

+ estimates not available (see figure 4-1 for explanation)

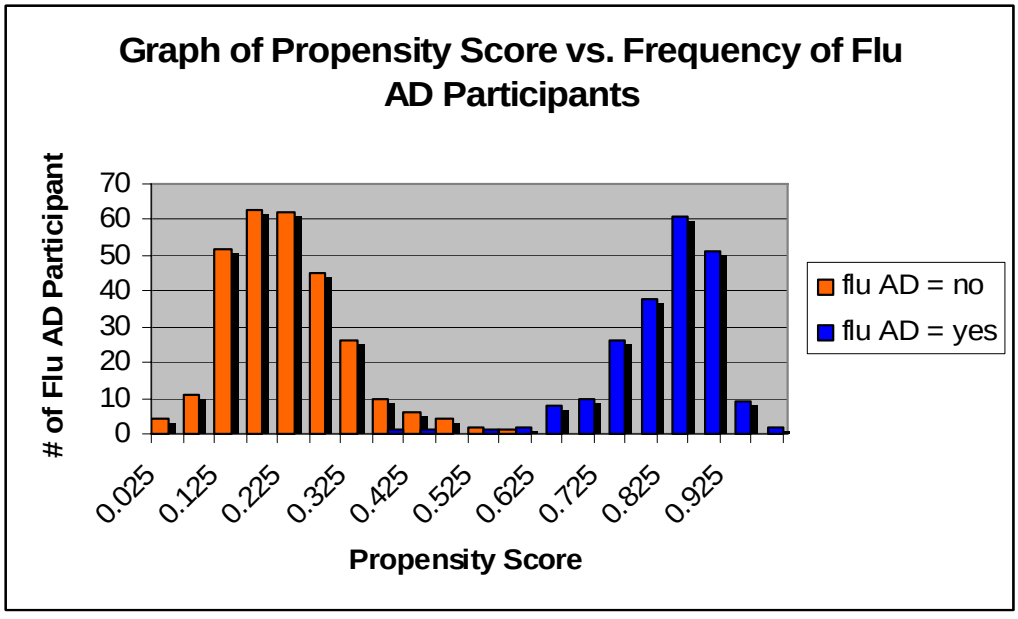

Figure 4-1. Frequency of Influenza AD Participants by Propensity Score.

The flu AD variable estimates were not obtained for the same reasons described in the regression on the PS method section. With the exception of the flu AD variable, the "greedy method" balanced all variables and associated interaction terms.

Table 4-8. "Greedy Matching" Method Analysis Results.

"Greedy Matching" Method									
	lsmean		Main Effect		Interaction Effect				
Variable	AD = 0 (n=104)	AD = 1 (n=104)	F	p	F	p	B	OR (exp B)	% bias reduction
% OA dx*	0.0722	0.0788	0.18	0.6712	0.01	0.9330			65.98
GP age	46.4	47.0	2.10	0.1492	1.99	0.1598			62.50
% elderly	0.1768	0.1876	0.39	0.5328	0.05	0.8299			21.74
Total # pt	1074	1008	1.25	0.2650	0.37	0.5410			-312.50
aver income*	26535	26300	0.23	0.6301	0.12	0.7308			87.34
BL rate	3.75	3.78	0.85	0.3573	1.25	0.2651			92.37
los rate*	0.0456	0.0535	1.04	0.3097	0.89	0.3471			81.59
population*	154497	132123	2.06	0.1527	1.14	0.2862			63.10
Sex			0.53	0.4654	1.26	0.2619	0.4266	1.5320	-1538.99
flu AD*			xx[+]	xx[+]	xx[+]	xx[+]	446.3283		
License			1.35	0.2459	1.04	0.3077	-0.6730	0.5102	-18.49
birth place			1.79	0.1809	1.31	0.2520	-0.6828	0.5052	-819.72
Average**									74.50

* variables that showed significant differences at the alpha = 0.05 level in the pre-PS model

** average % bias reduction for variables that were not significant in the pre-PS model (excluding flu AD)

[+] estimates not available

The "greedy matching" method resulted in a decrease in total sample size from 496 (sample size of the two previous methods) to 208. This represents a decrease in sample size of 58%. The eliminated GPs had PSs that were predominantly in the highest or lowest ranges of the distribution. The elimination of these GPs could affect the generalizability of the study since only the GPs who are in the midrange of the PS distribution would be left in the study.

Selection of a Preferred Propensity Score Method

The selection of a preferred PS method was carried out by measuring each of the three methods against the following two criteria;

- the resulting sample size after application of the PS method, and

- the PS method's ability to adjust for bias on the VOC.

A major disadvantage of the "greedy matching" method is the reduction in sample size resulting from the discarding of subjects that are not matched. In this case the sample size is reduced by 58% which possibly results in a loss of power to detect significance in the main effects of the outcome models and a loss of generalizability of the findings. Since the "greedy matching" method does not show advantages over the regression on the PS method in terms of adjusting for bias on the covariates it is considered less desirable than the regression on the PS method and will not be selected as the PS method for inclusion in the outcome models.

The regression on PS method was responsible for the greatest adjustment for bias between groups on all of the VOC (figure 4-2). The average reduction in bias for the regression on the PS method was 99% versus 82% for the quintile method.

With this dataset the regression on the PS method is preferred and it is the method that will be applied to the outcome analyses. It is important to note that the failure to adjust for bias on the flu AD variable still exists and as such the flu AD variable will be included in the outcome models.

Exploratory Analysis of the Propensity Score Methods Effect on Adjusting for Bias on Unmeasured Variables

The purpose of this exploratory analysis is to determine whether any one PS method is better at reducing bias on variables that are not included in the PS model and are, therefore, considered unmeasured.

The c-statistic is a measure of the model's ability to discriminate between groups. The c-statistic for the full model is 0.832 which can be interpreted as follows; if one randomly select one subject from each AD group the model will accurately predict the

group from which the subjects originated 83.2% of the time. With the exception of the models dealing with the exclusion of the flu AD variable, the c-statistic remains stable for all of the PS models. The range is from 0.830 to 0.835 (table 4-9).

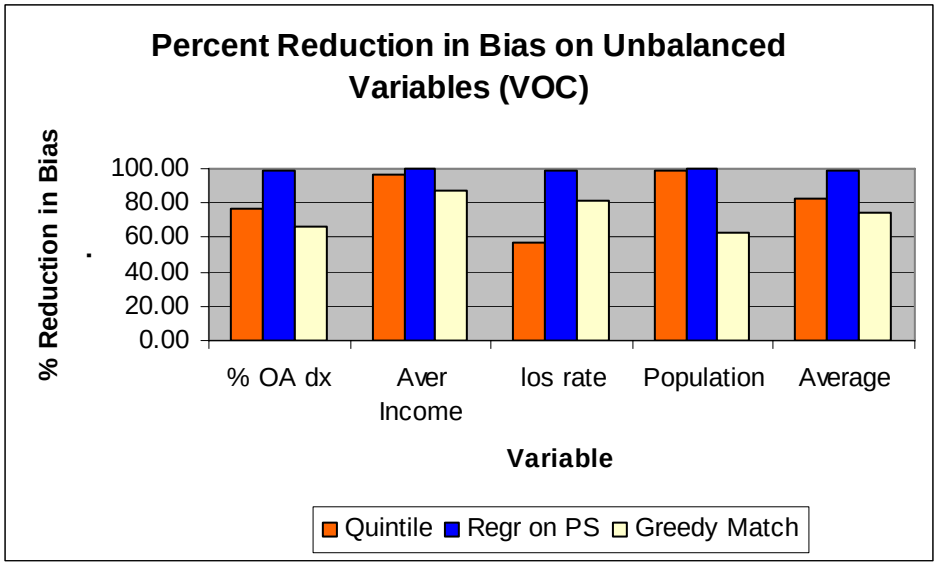

Figure 4-2. Comparison of PS methods Ability to Reduce Bias on VOC.

Table 4-9. Quintile Method Results for Excluded Variable Models.

	Quintile Method									
		lsmean		Main Effect		Interaction Effect				
Excluded Variable	c	AD = 0 (n=265)	AD = 1 (n=231)	F	p	F	p	B	OR (exp B)	% bias reduction
% OA dx*	0.833	0.0935	0.0708	4.92	0.0207	0.79	0.5298			-17.01
GP age	0.833	46.8	47.5	0.37	0.5426	0.82	0.5112			56.25
% elderly	0.831	0.1960	0.1754	3.10	0.079	1.75	0.1374			-49.28
total # pt	0.833	1085	1006	2.55	0.1111	2.86	0.0231			-393.75
aver income*	0.834	26777	26652	0.07	0.7862	1.39	0.2361			93.27
BL rate	0.834	3.6	3.7	0.11	0.7418	0.80	0.5274			73.68
los rate*	0.833	0.0471	0.0730	1.97	0.1615	0.81	0.5214			39.63
population*	0.832	153438	148662	0.08	0.7746	1.57	0.1805			92.12
sex	0.835			0.01	0.9217	0.25	0.6197	-0.0593	0.9424	-77.37
flu AD*	0.662			14.5	0.0001	1.09	0.2957	-2.0411	0.1299	6.38
license	0.830			0.83	0.3673	0.00	0.0896	-0.5001	0.6065	4.81
birth place	0.835			0.10	0.7528	0.02	0.8820	-0.1524	0.8586	-162.75
Average**										42.88

* variables that showed significant differences at the alpha = 0.05 level in the pre-PS model

** average % bias reduction for variables that showed significant differences in the pre-PS model

There are three c-statistics that are worth noting. The first is the c-statistic that is generated for the model when the flu AD variable is removed. It has been noted that there exists an almost complete separation of data for the flu AD variable on the PS so when the flu AD variable is excluded from the model the ability of the model to discriminate decreases from 0.834 to 0.662. The other two are the c-statistics associated with sex and birth place. These two variables have the distinction of being the most closely balanced variables in the pre-PS analysis (table 4-3) with p-values of 0.8663 and 0.7415 respectively. The PS model c-statistics when these variables are excluded is equal to 0.835 in both cases. This value is greater than the c-statistic for the full model thereby indicating that the inclusion of these variables in the PS model decreases the model's discriminative ability.

The complete results from the reduced PS models are reported in tables 4-9 through 4-11. The analysis of the reduced models effect's on balancing the VOC is summarized in figure 4-3. Figure 4-3 shows that no one PS method systematically reduces bias on unmeasured variables to a greater extent than the others. Regression on PS does, on average, reduce bias on the VOC to the greatest degree.

The summary of PS models effects (figure 4-3) shows that bias between groups on unmeasured variables can be reduced by PS methods. The correlation matrix between the VOC and the PS covariates was calculated and reported in table 4-12. Table 4-12 shows limited correlation (less than 0.30) between the VOC and the PS covariates in all cases except one. The one exception is the correlation between population (population of community where the GP practice is located) and aver income (average income for county where GP practice located) which was 0.91. The correlation between population

and aver income is associated with the higher reduction in bias for those variables when they are not included in the PS model.

Table 4-10. Regression on PS Results for Excluded Variable Models.

Regression on Propensity Score Method										
		lsmean		Main Effect		Interaction Effect				
Excluded Variable	c	AD = 0 (n=265)	AD = 1 (n=231)	F	p	F	p	B	OR (exp B)	% bias reduction
% OA dx*	0.833	0.0901	0.0736	0.74	0.3909	0.00	0.9540			14.95
GP age	0.833	46.5	47.2	1.14	0.2854	0.72	0.3974			56.25
% elderly	0.831	0.1918	0.1777	0.27	0.6065	0.04	0.8445			-2.17
total # pt	0.833	1061	1018	0.69	0.4080	0.15	0.6946			-168.75
aver income*	0.834	26567	26510	2.83	0.0929	3.65	0.0566			96.93
BL rate	0.834	3.7	3.7	0.57	0.4515	0.96	0.3286			100.00
Los rate*	0.833	0.0513	0.0684	0.19	0.6641	1.26	0.2618			60.14
population*	0.832	146919	139289	4.70	0.0307	5.12	0.0240			87.42
sex	0.835			0.06	0.7993	0.68	0.4108	0.1155	1.1224	-277.17
Flu AD*	0.662			3.16	0.0756	2.12	0.1456	-1.4258	0.2403	18.26
license	0.830			1.41	0.2348	4.59	0.0322	-0.5203	0.5943	1.87
birth place	0.835			0.09	0.7695	0.02	0.8911	-0.1117	0.8943	-96.45
Average**										55.54

* variables that showed significant differences at the alpha = 0.05 level in the pre-PS model

** average % bias reduction for variables that showed significant differences in the pre-PS model

Table 4-11. "Greedy Matching" Results for Excluded Variable Models.

"Greedy Matching" Method										
		lsmean		Main Effect		Interaction Effect				
Excluded Variable	ni (i=0,1)	AD = 0	AD = 1	F	p	F	p	B	OR (exp B)	% bias reduction
% OA dx*	106	0.0969	0.0707	4.36	0.0380	0.00	0.9677			-35.05
GP age	101	46.2	47.5	1.40	0.2381	0.59	0.4450			15.69
% elderly	105	0.1944	0.1801	0.41	0.5205	0.01	0.9365			-3.62
total # pt	103	1049	1005	0.61	0.4371	0.17	0.6676			-175.00
aver income*	105	26875	26123	1.02	0.3136	0.22	0.6419			59.48
BL rate	103	3.7	3.7	0.77	0.3806	1.01	0.3169			97.11
Los rate*	105	0.0352	0.0614	1.09	0.2967	0.06	0.8145			38.93
population*	104	154497	132123	2.06	0.1527	1.14	0.2862			63.10
sex	104			0.00	0.9674	0.24	0.6260	0.0119	1.0120	63.12
Flu AD*	190			2.00	0.1570	1.34	0.2471	-0.7168	0.4883	44.94
license	104			5.04	0.0248	9.09	0.0026	-0.6618	0.5159	-17.10
birth place	105			0.00	0.9787	0.28	0.5943	0.0069	1.0069	87.22
Average**										34.28

* variables that showed significant differences at the alpha = 0.05 level in the pre-PS model

** average % bias reduction for variables that showed significant differences in the pre-PS model

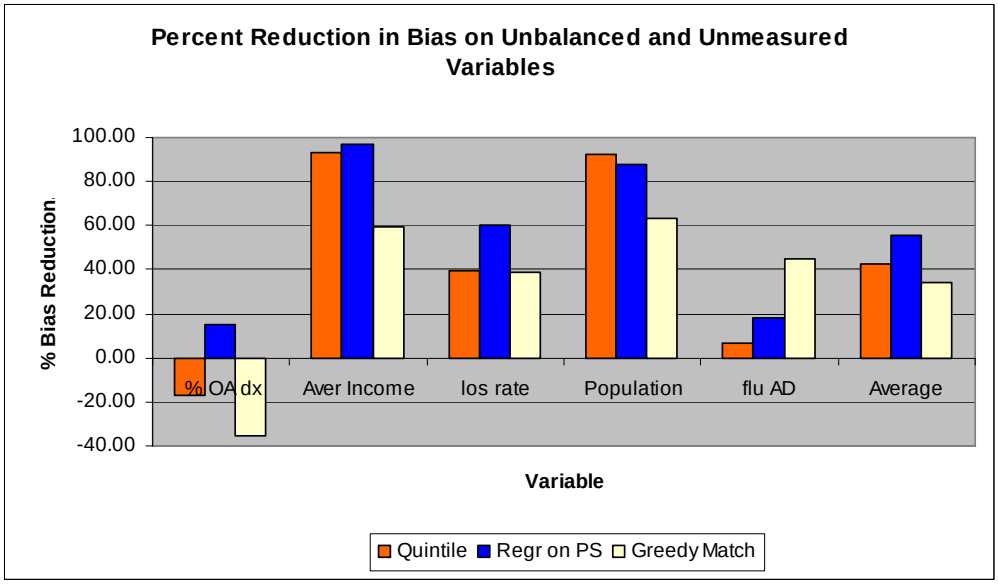

Figure 4-3. Summary of PS Models Effects on Reducing Bias on the VOC

Table 4-12. Correlation Matrix Between VOC and PS Covariates.

Correlation Between VOC and All PS Covariates					
Covariate	**VOC**				
	los rate	**% OA dx**	**population**	**aver income**	**flu AD**
BL rate	-0.06	-0.02	-0.22	-0.26	0.00
los rate	1.00	-0.01	-0.05	-0.05	0.03
% OA dx	-0.01	1.00	0.01	0.01	0.01
total # pt	-0.12	-0.10	0.00	0.02	0.02
% elderly	-0.02	0.26	0.09	0.09	-0.03
sex	-0.08	0.04	-0.13	-0.13	0.03
flu AD	0.03	0.01	-0.14	-0.17	1.00
population	-0.05	0.01	1.00	0.91	-0.14
GP age	-0.05	0.07	0.06	0.06	-0.09
aver income	-0.05	0.01	0.91	1.00	-0.17

The effect of the correlation between the PS and the VOC and the reduction in bias was tested. The correlation between the PS and the VOC was calculated and scatter plots were compiled to display the results graphically in figure 4-4.

The absolute values of the correlations ranged from 0.182 to 0.329. The absolute value of the correlations was plotted against the percent bias reductions on the VOC for each of the three PS methods (figure 4-5). The results from figure 4-5 show an overall

effect of increasing percent bias reduction with increasing absolute correlation between the PS and the excluded variable.

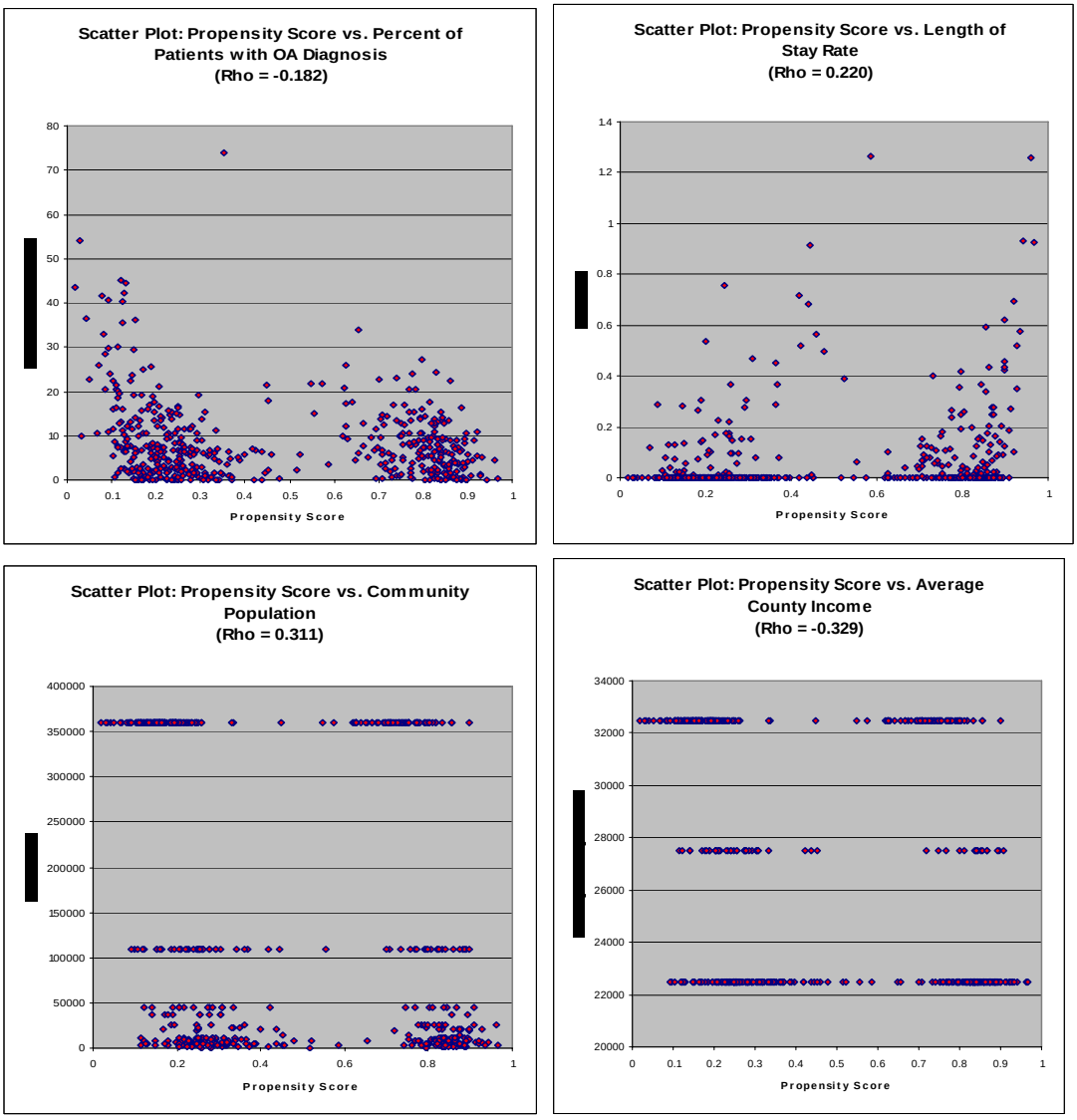

Figure 4-4. Scatterplots of Propensity Score Versus Unbalanced Variables

Step 3: Primary Outcome Analysis

Model Development

The analysis of the primary outcome, the effect of the OA AD intervention on the COX-2 utilization rates was carried out using a repeated measures model on longitudinal

data (PROC GENMOD[35]). There were six experimental time periods over which the outcomes measures were assessed (figure 3-2).

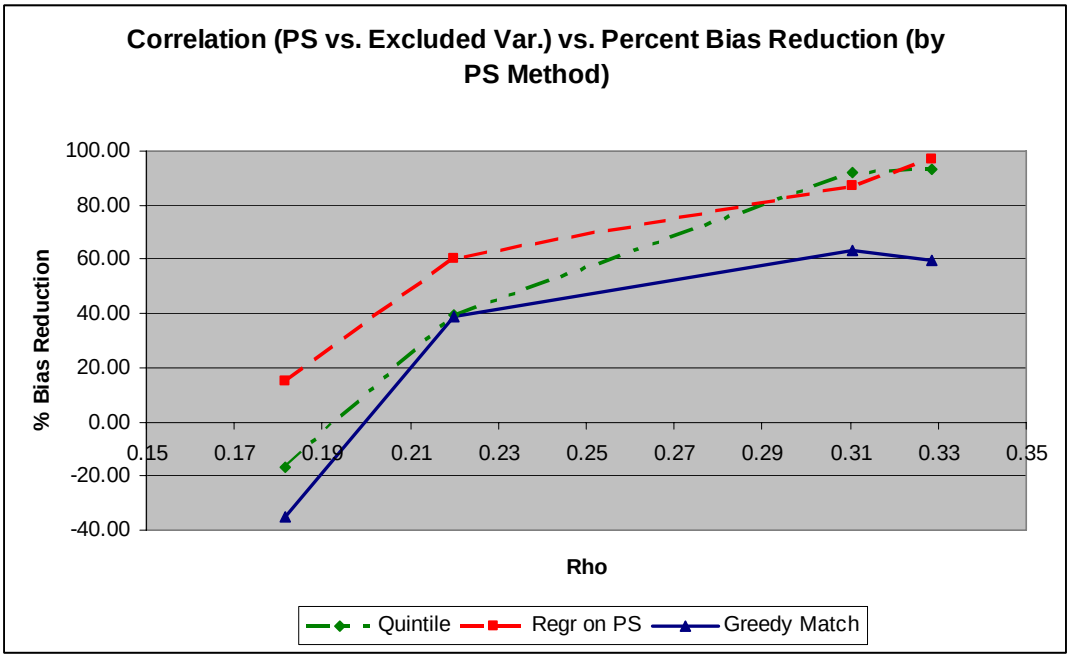

Figure 4-5. Line Graph Comparing Correlations and Percent Bias Reduction

The primary outcome measure, the change in COX-2 prescribing from baseline, was calculated for each physician by aggregating all of the COX-2 prescription claims for all of the elderly patients in the physician's panel and dividing by the number of elderly patients in the panel. The resulting rate, number of COX-2 DDDs per patient per physician was subtracted from the baseline prescribing rate to yield a measure of change in COX-2 prescribing.

The primary outcome model included the variables intervention participation (AD), the PS (pr), the time period in which the measurement took place (period), participation in a previous influenza AD service (flu AD), the baseline COX-2 prescribing rate (BL rate), and the number of elderly patients in the GP's panel (# elderly). The model is depicted in figure 4-6. The variables were included for the following reasons. The PS

variable represents the outcome from the PS analysis, the period variable controls for the longitudinal changes, the flu AD variable was not successfully balanced by the PS method, and the baseline COX-2 rate and number of elderly patients control for the GP's pre-intervention prescribing behavior and practice size respectively.

$$\mathbf{Y = \beta_0 + \beta_1(X_1) + \beta_2(X_2) + \beta_3(X_3) + \beta_4(X_4) + \beta_5(X_5) + \beta_6(X_6)}$$

Where;
Y = change in COX-2 utilization rate (periods 3 to 6 (post-intervention)),
X_1 = physician participation in the intervention (0 = no, 1 = yes),
X_2 = PS (range from 0 to 1),
X_3 = experimental time period (period = 3,4,5,6),
X_4 = physician participation in the influenza AD service (0 = no, 1 = yes),
X_5 = physician baseline COX-2 rate (DDD / patient, (period = 2)),
X_6 = number of patients in the GP's practice >65 years old

Figure 4-6. Primary Outcome Model

Between Group Results

The significance level of each variable from the primary outcome model (figure 4-6) is listed in table 4-13. The values of the coefficient estimates in GEE are not interpreted in the same manner as GLM models[37] and as such the values of the coefficient estimates are not reported in the results tables. A more in-depth discussion of the interpretation of GEE results is included in the discussions in chapter five.

The between groups effect of the intervention is interpreted from the value of the z statistic for the AD variable. The z value of 0.85 and associated p-value of 0.3976 indicates that the main intervention effect over the entire post-intervention period is not statistically significant.

The model in figure 4-6 was also used to determine between group differences in the pre-intervention time periods (period = 1, 2). The z statistics and associated p-values of each variable are listed in table 4-14. The pre-intervention results are interpreted in the

same manner as the post-intervention results. The z statistic and p-value for the AD variable are 0.88 and 0.3775 respectively. The p-value indicates that the groups are not significantly different on the outcome measure in the pre-intervention periods at the alpha = 0.05 level.

Table 4-13. Primary Outcome Model Results (Periods = 3,4,5,6).

Primary Outcome Model Results for Post-intervention Periods (COX-2 Prescribing Rates)		
Effect	Z	p-value
AD (AD = no)	0.85	0.3976
PS	0.84	0.4023
period	2.69	0.0072
flu AD (flu AD = no)	1.21	0.2255
BL rate	-10.68	<0.0001
# elderly	1.64	0.1017

Table 4-14. Primary Outcome Model Results (Periods = 1,2).

Primary Outcome Model Results for Pre-intervention Periods (COX-2 Prescribing Rates)		
Effect	Z	p-value
AD (AD = no)	0.88	0.3775
PS	0.31	0.7588
period	0.38	0.7018
flu AD (flu AD = no)	0.23	0.8170
BL rate	-14.45	<0.0001
# elderly	-0.48	0.6313

Table 4-15 depicts the least square means for the two groups (AD = yes and AD = no) for each of the six experimental time periods. The least square mean values are also presented in a graph in figure 4-7.

Table 4-16 depicts the unadjusted means and standard deviations for the two groups (AD = yes and AD = no) for each of the six experimental time periods. A positive value indicates that the prescribing rate has increased from the baseline rate by the amount indicated and a negative value indicates a decrease in the prescribing rate from baseline. The unadjusted mean values are also presented in a graph in figure 4-8.

Table 4-15. Least Square Means for Change in COX-2 Rates by Group (DDDs/patient).

AD group	period					
	1	2	3	4	5	6
0	0.0516	0	0.2275	-0.1778	0.2471	0.4178
1	-0.2136	0	-0.5315	-0.0018	0.261	0.1457

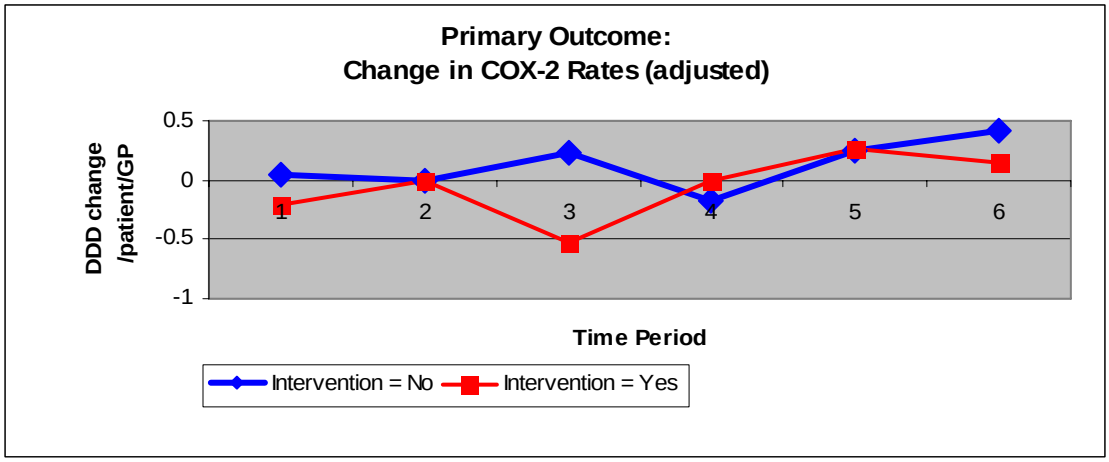

Figure 4-7. Least Square Means for Change in COX-2 Rates by Group

Table 4-16. Unadjusted Means for Change in COX-2 Rates by Group (DDDs/patient).

Period	AD group			
	0		1	
	Mean	Std Dev	Mean	Std Dev
1	0.1587	3.4287	-0.3026	3.0709
2	0.0000	0.0000	0.0000	0.0000
3	0.3396	3.9059	-0.5321	3.4410
4	0.0236	3.6700	-0.1257	3.7358
5	0.5266	3.7285	-0.0008	3.9072
6	0.6454	3.8566	0.1281	3.9248

Within Group (Longitudinal) Results

The within group models were the same as the between group model in figure 4-6 except that the AD group variable is replaced by a prepost variable which measures significant within group differences between change in COX-2 rates pre-intervention and post-intervention. The model is run two times; once including only the intervention group and once including only the control group. The z statistic value and associated

significance level (p-value) of each variable are listed in table 4-17 for the intervention group and table 4-18 for the control group.

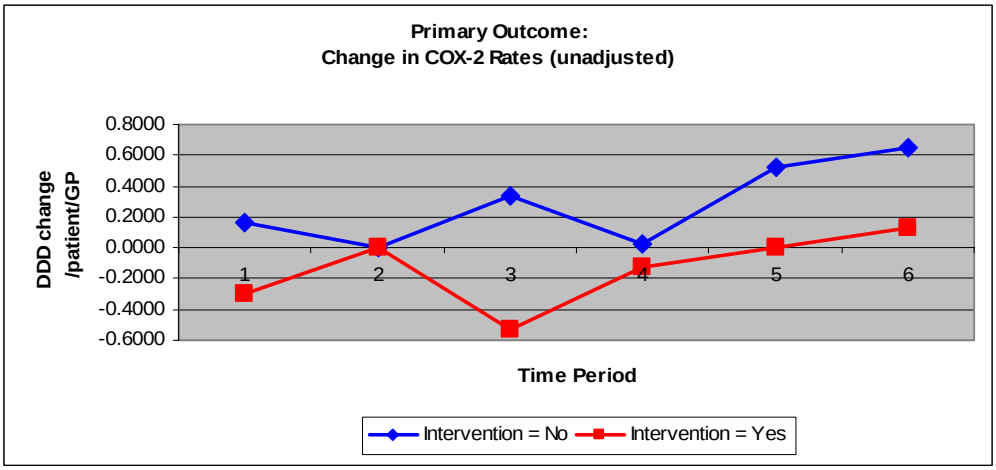

Figure 4-8. Unadjusted Means for Change in COX-2 Rates by Group

The within group effect of the intervention is interpreted from the values of the z statistic and significance level of the prepost variable. For the intervention group, the z and p-values of -2.34 and 0.0191 respectively indicates that the within group effect is significant at the alpha = 0.05 level. For the control group, the z statistic and p-value of --0.22 and 0.8273 respectively indicates that the within group effect is not significant at the alpha = 0.05 level.

Table 4-17. Primary Outcome Model Results (AD = yes).

Primary Outcome Results for the Intervention Group (COX-2 Prescribing Rates)		
Effect	Z	p-value
PS	0.04	0.9708
period	2.82	0.0049
prepost	-2.34	0.0191
flu AD (flu AD = no)	0.49	0.6217
BL rate	-9.74	<0.0001
# elderly	0.63	0.5271

Step 4: Secondary Outcome Analyses

Misoprostol Utilization Rates

Model development

The analysis of the secondary outcome, the effect of the OA AD intervention on the misoprostol utilization rate was carried out using the same methods as the primary outcome analysis with the data for misoprostol utilization substituted for the COX-2 utilization data (figure 4-9).

Table 4-18. Primary Outcome Model Results (AD = no).

Primary Outcome Results for the Control Group (COX-2 Prescribing Rates)		
Effect	Z	p-value
PS	1.32	0.1881
period	1.31	0.1910
prepost	-0.22	0.8273
flu AD (flu AD = no)	0.95	0.3412
BL rate	-8.68	<0.0001
# elderly	1.10	0.2727

$$\mathbf{Y = \beta_0 + \beta_1(X_1) + \beta_2(X_2) + \beta_3(X_3) + \beta_4(X_4) + \beta_5(X_5) + \beta_6(X_6)}$$

Where;
Y = change in misoprostol utilization rate (periods 3 to 6 (post-intervention)),
X_1 = physician participation in the intervention (0 = no, 1 = yes),
X_2 = PS (range from 0 to 1),
X_3 = experimental time period (period = 3,4,5,6),
X_4 = physician participation in the influenza AD service (0 = no, 1 = yes),
X_5 = physician baseline misoprostol rate (DDD / patient, (period = 2)),
X_6 = number of patients in the GP's practice >65 years old

Figure 4-9. Secondary Outcome Model for Misoprostol Utilization

Between group results

The z statistic and the significance level (p-value) of each variable from the secondary misoprostol outcome model (figure 4-9) are listed in table 4-19.

The between group effect of the intervention is interpreted from the z statistics and associated p-value of the AD variable. The z statistic and p-value of -0.87 and 0.3866 respectively indicate that the effect is not significant at the alpha = 0.05 level.

Table 4-19. Secondary Misoprostol Outcome Model Results (Periods = 3,4,5,6).

Secondary Outcome Model Results for Post-intervention Periods 3 to 6
(Change in Misoprostol Prescribing Rates)

Effect	Z	p-value
AD (ad = 0)	-0.87	0.3866
PS	-0.61	0.5412
period	0.96	0.3359
flu AD (flu AD = 0)	-0.53	0.5943
BL rate	-6.31	<0.0001
# elderly	-0.24	0.8091

The model in figure 4-9 was used to determine intervention effects on each post intervention time period. None of the post intervention (analyzed individually) showed significant between group differences at the alpha = 0.05 level.

The model in figure 4-9 was used to determine between group differences in the pre-intervention time periods (period = 1, 2). The z statistic and associated significance level of each variable is listed in table 4-20. The results are interpreted in the same manner as the post-intervention results. The z statistic and p-value for the AD variable are -0.22 and 0.8269 respectively. The p-value indicates that the groups are not significantly different in the pre-intervention periods at the alpha = 0.05 level.

Table 4-20. Secondary Misoprostol Outcome Model Results (Periods = 1,2).

Secondary Outcome Model Results for Post-intervention Periods 1 and 2
(Change in Misoprostol Prescribing Rates)

Effect	Z	p-value
AD (ad = 0)	-0.22	0.8269
PS	1.20	0.2308
period	0.28	0.7758
flu AD (flu AD = 0)	1.18	0.2396
BL rate	-4.55	<0.0001
# elderly	0.58	0.5612

Table 4-21 depicts the least square means for the two groups (AD = yes and AD = no) for each of the six experimental time periods. The least square mean values are also presented in a graph in figure 4-10.

Table 4-21. Least Square Means for Change in Misoprostol Rate by Group (DDDs/patient).

Secondary Outcome (Misoprostol) Least Square Means by AD Group						
AD group	Period					
	1	2	3	4	5	6
0	-0.0191	0.0000	0.0172	0.0516	0.0390	0.0388
1	-0.0127	0.0000	0.0383	0.0743	0.0604	0.0652

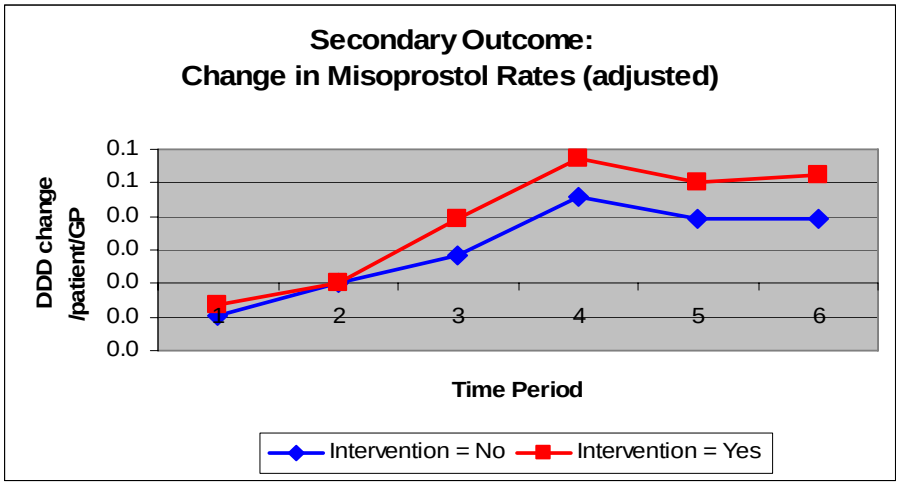

Figure 4-10. Least Square Means for Change in Misoprostol Rates by Group.

Table 4-22 depicts the unadjusted means and standard deviations for the two groups (AD = yes and AD = no) for each of the six experimental time periods. The unadjusted mean values are also presented in a graph in figure 4-11.

Within group (longitudinal) results

The within group model was the same as the between group model in figure 4-9 except that the AD group variable is replaced by a prepost variable which measures within group differences between change in misoprostol rates pre-intervention and post-intervention. The model is run two times; once including only the intervention group and

once including only the control group. The z statistic and the associated significance level (p-value) of each variable are listed in table 4-23 for the intervention group and table 4-24 for the control group.

Table 4-22. Unadjusted Means and Standard Deviations for Change in Misoprostol Rate by Group (DDDs/patient).

Period	AD group			
	0		1	
	Mean	Std Dev	Mean	Std Dev
1	-0.0153	0.2877	0.0096	0.2969
2	0.0000	0.0000	0.0000	0.0000
3	0.0032	0.2693	0.0437	0.2581
4	0.0586	0.3289	0.0570	0.3120
5	0.0484	0.3750	0.0497	0.3472
6	0.0235	0.3782	0.0850	0.4005

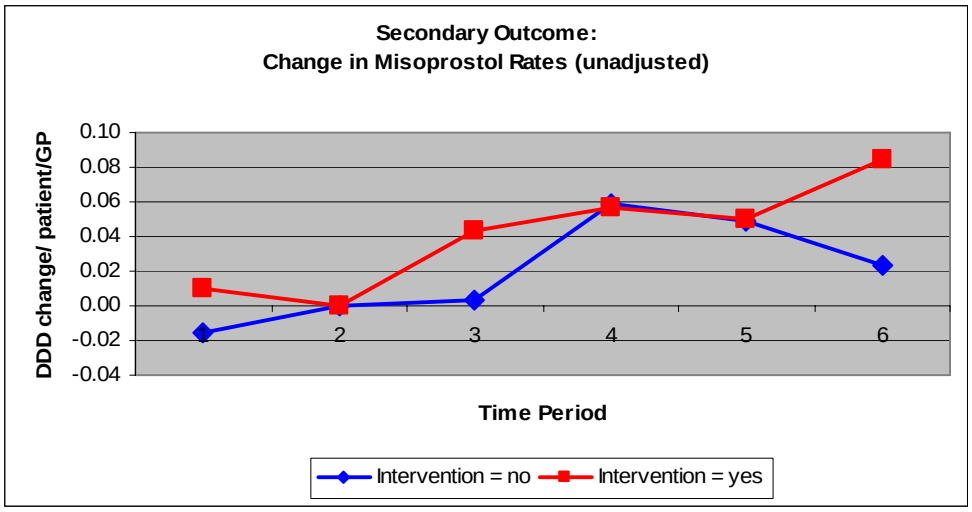

Figure 4-11. Unadjusted Means for Change in Misoprostol Rates by Group.

Table 4-23. Secondary Misoprostol Outcome Model Results (AD = yes).

Secondary Outcome Model Results for All Periods (Intervention Group) (Change in Misoprostol Prescribing Rates)		
Effect	Z	p-value
PS	-0.58	0.5594
period	1.65	0.0990
prepost	0.25	0.8075
flu AD (flu AD = 0)	-0.30	0.7612
BL rate	1.23	0.2195
# elderly	0.55	0.5802

Table 4-24. Secondary Misoprostol Outcome Model Results (AD = no).

Secondary Outcome Model Results for All Periods (Control Group) (Change in Misoprostol Prescribing Rates)		
Effect	Z	p-value
PS	-0.01	0.9921
period	0.75	0.4523
prepost	1.00	0.3176
flu AD (flu AD = 0)	-0.27	0.7888
BL rate	4.69	<0.0001
# elderly	1.59	0.1109

The within group effect of the intervention is interpreted from the values of the z statistic and associated p-value of the prepost variable. For the intervention and control groups, z statistics and the p-values of 0.25, 0.8075 and 1.00, 0.3176 respectively indicates that the within group effect is not statistically significant for both groups at the alpha = 0.05 level.

PPI Utilization Rates

Model development

The analysis of the secondary outcome, the effect of the OA AD intervention on the PPI utilization rates was carried out using the same methods as the primary outcome analysis with the data for PPI utilization substituted for the COX-2 utilization data (figure 4-12).

$$\mathbf{Y = \beta_0 + \beta_1(X_1) + \beta_2(X_2) + \beta_3(X_3) + \beta_4(X_4) + \beta_5(X_5) + \beta_6(X_6)}$$

Where;
Y = change in PPI utilization rate (periods 3 to 6 (post-intervention)),
X_1 = physician participation in the intervention (0 = no, 1 = yes),
X_2 = PS (range from 0 to 1),
X_3 = experimental time period (period = 3,4,5,6),
X_4 = physician participation in the influenza AD service (0 = no, 1 = yes),
X_5 = physician baseline PPI rate (DDD / patient, (period = 2)),
X_6 = number of patients in the GP's practice >65 years old

Figure 4-12. Secondary PPI Outcome Model

Between group results

The z statistic and the significance level (p-value) of each variable from the secondary PPI outcome model (figure 4-12) are listed in table 4-25.

Table 4-25. Secondary PPI Outcome Model Results (Periods = 3,4,5,6).

Secondary Outcome Model Results for Post-intervention Periods 3 to 6 (Change in PPI Prescribing Rates)		
Effect	Z	p-value
AD (ad = 0)	-0.27	0.7906
PS	1.09	0.2755
period	1.43	0.1519
flu AD (flu AD = 0)	1.45	0.1478
BL rate	-2.92	0.0035
# elderly	-1.74	0.0813

The between group effect of the intervention is interpreted from the z statistics and associated p-value of the AD variable. The z statistic and p-value of -0.27 and 0.7906 respectively indicate that the effect is not significant at the alpha = 0.05 level.

The model (figure 4-12) was used to determine intervention effects on each post intervention time period. None of the post intervention (analyzed individually) showed significant between group differences at the alpha = 0.05 level.

The model in figure 4-12 was used to determine between group differences in the pre-intervention time periods (period = 1, 2). The z statistic and associated significance level of each variable is listed in table 4-26. The results are interpreted in the same manner as the post-intervention results. The z statistic and p-value for the AD variable are 0.13 and 0.8989 respectively. The p-value indicates that the groups are not significantly different in the pre-intervention periods at the alpha = 0.05 level.

Table 4-27 depicts the least square means for the two groups (AD = yes and AD = no) for each of the six experimental time periods. The least square mean values are also presented in a graph in figure 4-13.

Table 4-26. Secondary PPI Outcome Model Results (Periods = 1,2).

Secondary Outcome Model Results for Post-intervention Periods 1 and 2 (Change in PPI Prescribing Rates)		
Effect	Z	p-value
AD (ad = 0)	0.13	0.8989
PS	1.10	0.2726
period	0.14	0.8911
flu AD (flu AD = 0)	1.08	0.2818
BL rate	-5.61	<0.0001
# elderly	0.04	0.9700

Table 4-27. Least Square Means for Change in PPI Rates by Group (DDDs/patient).

Secondary Outcome (PPI) Least Square Means by AD Group						
AD group	period					
	1	2	3	4	5	6
0	-0.0388	0	0.3194	0.5271	0.507	0.4842
1	-0.0553	0	0.3675	0.5513	0.555	0.4982

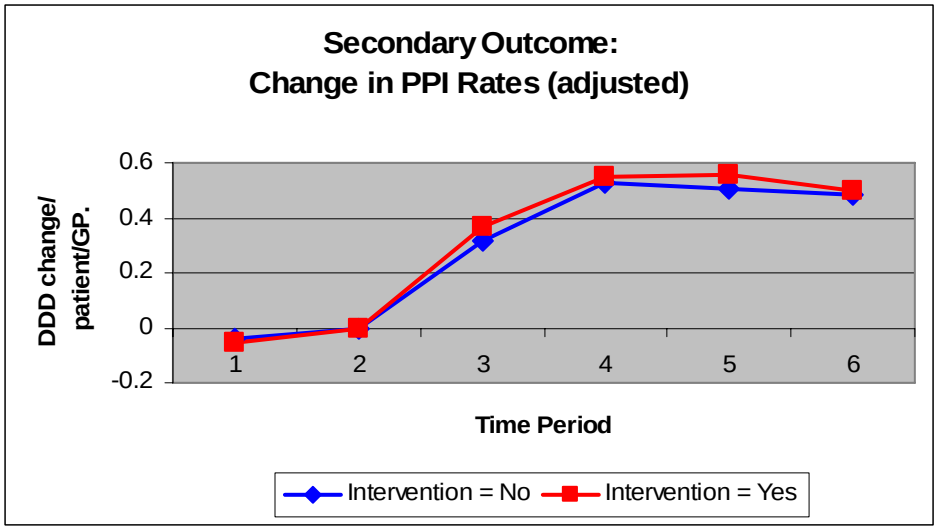

Figure 4-13. Least Square Means for Change in PPI Rates by Group.

Table 4-28 depicts the unadjusted means and standard deviations for the two groups (AD = yes and AD = no) for each of the six experimental time periods. The unadjusted mean values are also presented in a graph in figure 4-14.

Within group (longitudinal) results

The within group model was the same as the between group model in figure 4-12 except that the AD group variable is replaced by a prepost variable which measures

within group differences between change in PPI rates pre-intervention and post-intervention. The model is run two times; once including only the intervention group and once including only the control group. The z statistic and the associated significance level (p-value) of each variable are listed in table 4-29 for the intervention group and table 4-30 for the control group.

Table 4-28. Unadjusted Means for Change in PPI Rate by Group (DDDs/patient).

Period	OA group			
	0		1	
	Mean	Std Dev	Mean	Std Dev
1	-0.0203	1.4843	0.0041	1.4211
2	0.0000	0.0000	0.0000	0.0000
3	0.3932	1.3733	0.3372	1.5783
4	0.6307	1.6214	0.5810	1.6686
5	0.5844	1.6238	0.6182	1.7077
6	0.5575	1.7553	0.5490	1.6955

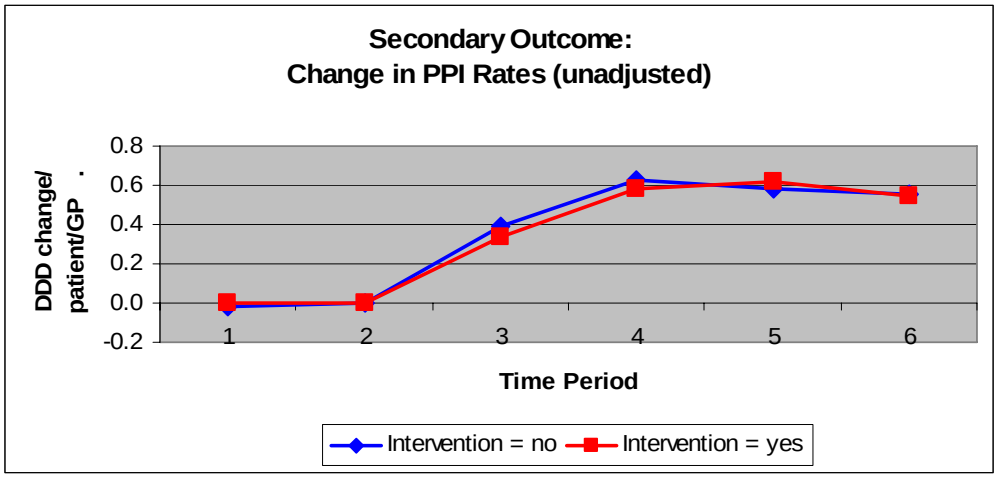

Figure 4-14. Unadjusted Means for Change in PPI Rates by Group.

The within group effect of the intervention is interpreted from the values of the z statistic and associated p-value of the prepost variable. For the intervention and control groups, the z statistics (p-values) of -2.59 (0.0097) and -4.22 (<0.0001) respectively indicates that the within group effect is statistically significant for both groups at the alpha = 0.05 level and both changes are in the direction of increased utilization.

Table 4-29. Secondary PPI Outcome Model Results (AD = yes).

Secondary Outcome Model Results for All Periods (Intervention Group) (Change in PPI Prescribing Rates)		
Effect	Z	p-value
PS	1.41	0.1596
period	0.69	0.4873
prepost	-2.59	0.0097
flu AD (flu AD = 0)	1.37	0.1717
BL rate	-3.63	0.0003
# elderly	-0.40	0.6879

Table 4-30. Secondary PPI Outcome Model Results (AD = no).

Secondary Outcome Model Results for All Periods (Control Group) (Change in PPI Prescribing Rates)		
Effect	Z	p-value
PS	0.67	0.5012
period	-0.02	0.9877
prepost	-4.22	<0.0001
flu AD (flu AD = 0)	1.16	0.2450
BL rate	-3.32	0.0009
# elderly	-1.61	0.1065

H2A Utilization Rates

Model development

The analysis of the secondary outcome, the effect of the OA AD intervention on the H2A utilization rates was carried out using the same methods as the primary outcome analysis with the data for H2A utilization substituted for the COX-2 utilization data (figure 4-15).

Between group results

The z statistic and the significance level (p-value) of each variable from the secondary H2A outcome model (figure 4-15) are listed in table 4-31.

The between group effect of the intervention is interpreted from the z statistics and associated p-value of the AD variable. The z statistic and p-value of 0.05 and 0.9619 respectively indicate that the effect is not significant at the alpha = 0.05 level.

$$\mathbf{Y = \beta_0 + \beta_1(X_1) + \beta_2(X_2) + \beta_3(X_3) + \beta_4(X_4) + \beta_5(X_5) + \beta_6(X_6)}$$

Where;
Y = change in H2A utilization rate (periods 3 to 6 (post-intervention)),
X_1 = physician participation in the intervention (0 = no, 1 = yes),
X_2 = PS (range from 0 to 1),
X_3 = experimental time period (period = 3,4,5,6),
X_4 = physician participation in the influenza AD service (0 = no, 1 = yes),
X_5 = physician baseline H2A rate (DDD / patient, (period = 2)),
X_6 = number of patients in the GP's practice >65 years old

Figure 4-15. Secondary Outcome Model for H2A Utilization

Table 4-31. Secondary H2A Outcome Model Results (Periods = 3,4,5,6).

Secondary Outcome Model Results for Post-intervention Periods 3 to 6 (Change in H2A Prescribing Rates)		
Effect	Z	p-value
AD (ad = 0)	0.05	0.9619
PS	1.18	0.2381
period	-7.29	<0.0001
flu AD (flu AD = 0)	1.12	0.2642
BL rate	-7.31	<0.0001
# elderly	1.77	0.0766

The model (figure 4-15) was used to determine intervention effects on each post intervention time period. None of the post intervention (analyzed individually) showed significant between group differences at the alpha = 0.05 level.

The model in figure 4-15 was used to determine between group differences in the pre-intervention time periods (period = 1, 2). The z statistic and associated significance level of each variable is listed in table 4-32. The results are interpreted in the same manner as the post-intervention results. The z statistic and p-value for the AD variable are 1.09 and 0.2764 respectively. The p-value indicates that the groups are not significantly different in the pre-intervention periods at the alpha = 0.05 level.

Table 4-33 depicts the least square means for the two groups (AD = yes and AD = no) for each of the six experimental time periods. The least square mean values are also presented in a graph in figure 4-16.

Table 4-32. Secondary H2A Outcome Model Results (Periods = 1,2).

Secondary Outcome Model Results for Post-intervention Periods 1 and 2 (Change in H2A Prescribing Rates)		
Effect	Z	p-value
AD (ad = 0)	1.09	0.2764
PS	0.98	0.3293
Period	-2.74	0.0062
flu AD (flu AD = 0)	0.64	0.5230
BL rate	-5.75	<0.0001
# elderly	1.49	0.1368

Table 4-33. Least Square Means for Change in H2A Rate by Group (DDDs/patient).

Secondary Outcome (H2A) Least Square Means by AD Group						
AD group	period					
	1	2	3	4	5	6
0	0.3007	0	0.1833	0.0114	-0.1433	-0.5532
1	0.0881	0	0.0157	0.1413	0.0867	-0.6818

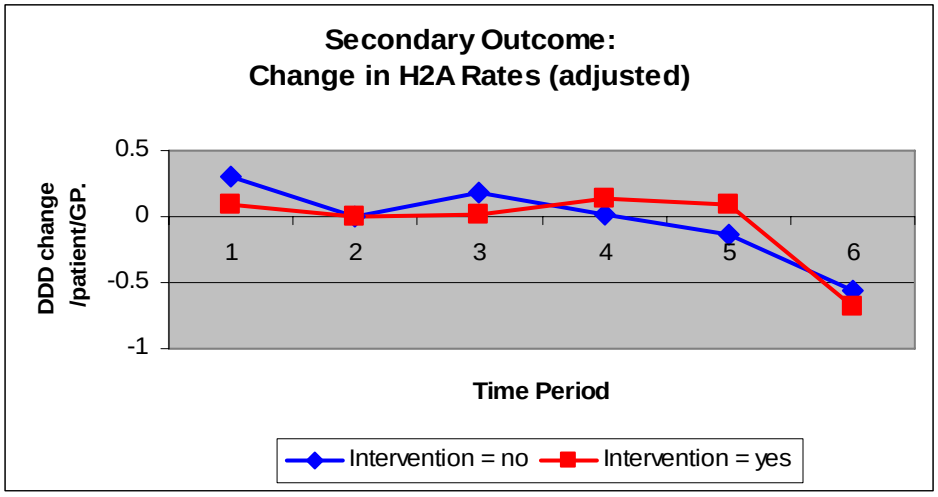

Figure 4-16. Least Square Means for Change in H2A Rates by Group.

Table 4-34 depicts the unadjusted means and standard deviations for the two groups (AD = yes and AD = no) for each of the six experimental time periods. The unadjusted mean values are also presented in a graph in figure 4-17.

Table 4-34. Unadjusted Means for Change in H2A Rate by Group (DDDs/patient).

Period	AD group			
	0		1	
	Mean	Std Dev	Mean	Std Dev
1	0.2440	1.7956	0.2117	1.9819
2	0.0000	0.0000	0.0000	0.0000
3	0.1655	1.9023	0.0783	2.0181
4	0.0607	1.8949	0.2191	2.2941
5	-0.2484	2.4465	0.2929	2.4570
6	-0.5101	2.5878	-0.5380	2.5854

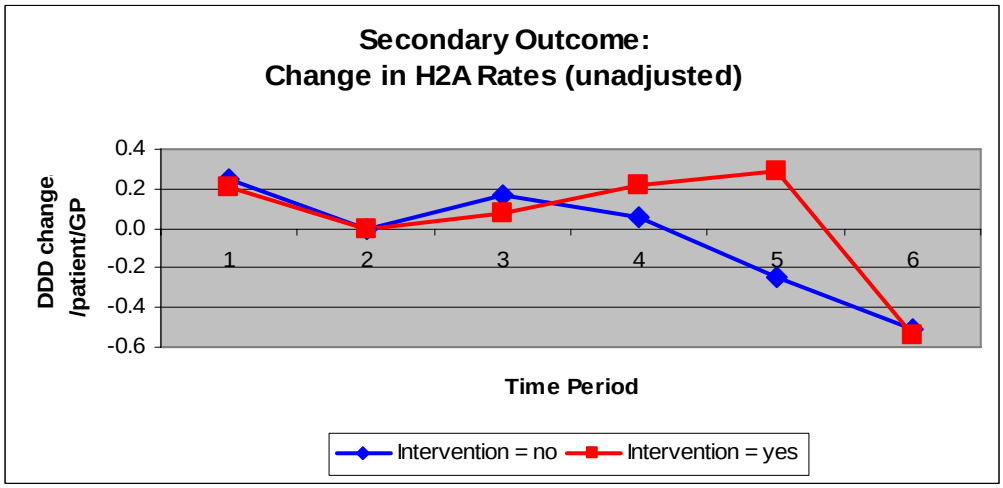

Figure 4-17. Unadjusted Means for Change in H2A Rates by Group.

Within group (longitudinal) results

The within group model was the same as the between group model in figure 4-15 except that the AD group variable is replaced by a prepost variable which measures within group differences between change in H2A rates pre-intervention and post-intervention. The model is run two times; once including only the intervention group and once including only the control group. The z statistic and the associated significance level (p-value) of each variable are listed in table 4-35 for the intervention group and table 4-36 for the control group.

The within group effect of the intervention is interpreted from the values of the z statistic and associated p-value of the prepost variable. For the intervention and control

groups, the z statistics (p-values) of -5.56 (<0.0001) and -4.06 (<0.0001) respectively indicates that the within group effect is statistically significant for both groups at the alpha = 0.05 level and both changes are in the direction of decreased utilization.

Table 4-35. Secondary H2A Outcome Model Results (AD = yes).

Secondary Outcome Model Results for All Periods (Intervention Group) (Change in H2A Prescribing Rates)		
Effect	Z	p-value
PS	1.70	0.0897
period	-6.59	<0.0001
prepost	-5.56	<0.0001
flu AD (flu AD = 0)	1.53	0.1262
BL rate	-7.48	<0.0001
# elderly	2.80	0.0051

Table 4-36. Secondary H2A Outcome Model Results (AD = no).

Secondary Outcome Model Results for All Periods (Control Group) (Change in H2A Prescribing Rates)		
Effect	Z	p-value
PS	-0.79	0.4282
period	-4.06	<0.0001
prepost	-2.33	0.0201
flu AD (flu AD = 0)	-0.78	0.4366
BL rate	3.43	0.0006
# elderly	-1.64	0.1003

GP Office Visit Rates

Model development

The analysis of the secondary outcome, the effect of the OA AD intervention on GP office visit rates was carried out using the same methods as the primary outcome analysis with the data for GP office visit rates substituted for the COX-2 utilization data (figure 4-18).

Between group results

The z statistic and the significance level (p-value) of each variable from the secondary GP office visit outcome model (figure 4-18) are listed in table 4-37.

$$Y = \beta_0 + \beta_1(X_1) + \beta_2(X_2) + \beta_3(X_3) + \beta_4(X_4) + \beta_5(X_5) + \beta_6(X_6)$$

Where;
Y = change in GP visit rates (periods 3 to 6 (post-intervention)),
X_1 = physician participation in the intervention (0 = no, 1 = yes),
X_2 = PS (range from 0 to 1),
X_3 = experimental time period (period = 3,4,5,6),
X_4 = physician participation in the influenza AD service (0 = no, 1 = yes),
X_5 = physician baseline GP visit rate rate (visits / patient, (period = 2)),
X_6 = number of patients in the GP's practice >65 years old

Figure 4-18. Secondary Outcome Model for GP Office Visits

Table 4-37. Secondary GP Office Visit Model Results (Periods = 3,4,5,6).

Secondary Outcome Model Results for Post-intervention Periods 3 to 6 (Change in GP Office Visit Rates)

Effect	Z	p-value
AD (ad = 0)	1.06	0.2888
PS	0.74	0.4587
period	-9.26	<0.0001
Flu AD (flu AD = 0)	0.02	0.9815
BL rate	-1.97	0.0487
# elderly	1.26	0.2077

The between group effect of the intervention is interpreted from the z statistics and associated p-value of the OA AD variable. The z statistic and p-value of 1.06 and 0.2888 respectively indicate that the effect is not significant at the alpha = 0.05 level.

The model (figure 4-18) was used to determine intervention effects on each post intervention time period. Only the period from 91 to 180 days (period four) following the intervention showed significant difference between groups at the alpha = 0.05 level. The z-statistic and p-value associated with the intervention effect are -2.20 and 0.0275 respectively (95% CI -0.7926, -0.0464). In this case, where the analysis only includes one time period, the interpretation of the coefficient estimate is similar to traditional GLM methods. That is, the coefficient estimate of -0.4195 (AD = no) is interpreted as the non-intervention group having measures of average change rate 0.4195 fewer

visits/patient/GP than the intervention group (equal values for the groups is hypothesized).

The model in figure 4-18 was used to determine between group differences in the pre-intervention time periods (period = 1, 2). The z statistic and associated significance level of each variable is listed in table 4-38. The results are interpreted in the same manner as the post-intervention results. The z statistic and p-value for the OA AD variable are 0.37 and 0.7097 respectively. The p-value indicates that the groups are not significantly different in the pre-intervention periods at the alpha = 0.05 level.

Table 4-38. Secondary GP Office Visit Outcome Model Results (Periods = 1,2).

Secondary Outcome Model Results for Post-intervention Periods 1 and 2 (Change in GP Office Visit Rates)		
Effect	Z	p-value
AD (ad = 0)	0.37	0.7097
PS	1.16	0.2457
Period	-0.08	0.9390
Flu AD (flu AD = 0)	1.19	0.2341
BL rate	-7.17	<0.0001
# elderly	2.13	0.0332

Table 4-39 depicts the least square means for the two groups (AD = yes and AD = no) for each of the six experimental time periods. The least square mean values are also presented in a graph in figure 4-19.

Table 4-39. Least Square Means for Change in GP Office Visit Rate by Group (visits/patient).

Secondary Outcome (GP Visits) Least Square Means by AD Group						
AD group	Period					
	1	2	3	4	5	6
0	-0.0182	0	0.4652	0.3882	-0.3346	-0.4341
1	-0.0563	0	0.3813	0.79	-0.0201	0.0069

Table 4-40 depicts the unadjusted means and standard deviations for the two groups (AD = yes and AD = no) for each of the six experimental time periods. The unadjusted mean values are also presented in a graph in figure 4-20.

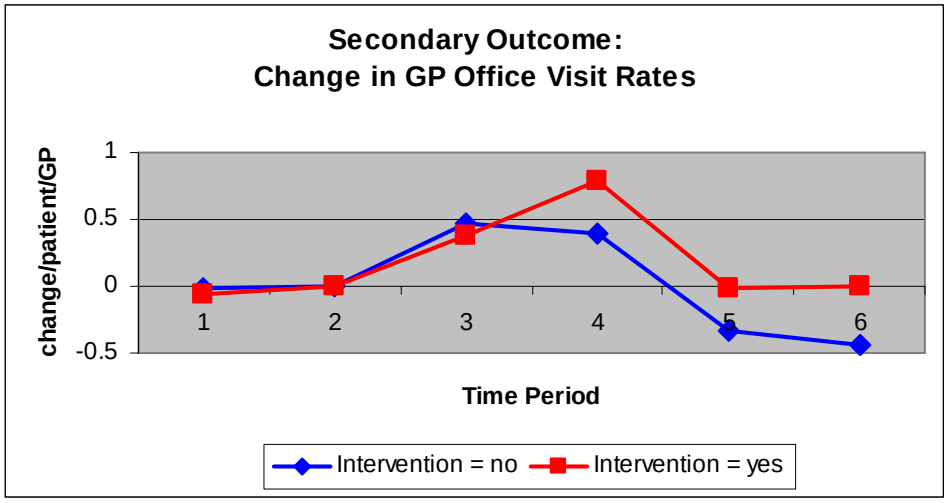

Figure 4-19. Least Square Means for Change in GP Office Visit Rates by Group.

Table 4-40. Unadjusted Means and Standard Deviations for Change in GP Office Visit Rate by Group (visits/patient).

Period	AD group			
	0		1	
	Mean	Std Dev	Mean	Std Dev
1	0.0057	0.5235	-0.0026	0.5586
2	0.0000	0.0000	0.0000	0.0000
3	0.5282	0.9720	0.3191	0.9419
4	0.2597	0.8025	0.6388	1.3687
5	-0.2269	0.7578	-0.0570	1.4364
6	-0.2742	0.7616	-0.0290	1.7479

Within group (longitudinal) results

The within group model was the same as the between group model in figure 4-18 except that the AD group variable is replaced by a prepost variable which measures within group differences between change in GP office visit rates pre-intervention and post-intervention. The model is run two times; once including only the intervention group and once including only the control group. The z statistic and the associated

significance level (p-value) of each variable are listed in table 4-41 for the intervention group and table 4-42 for the control group.

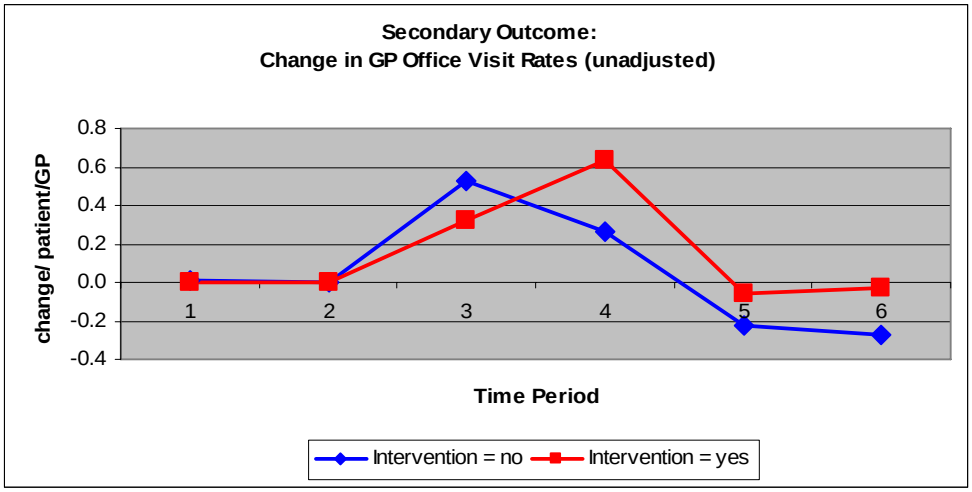

Figure 4-20. Unadjusted Means for Change in GP Office Visit Rates by Group.

Table 4-41. Secondary GP Office Visit Outcome Model Results (AD = yes).

Secondary Outcome Model Results for All Periods (Intervention Group) (Change in GP Office Visit Rates)		
Effect	Z	p-value
PS	-1.56	0.1199
Period	-10.95	<0.0001
Prepost	-17.54	<0.0001
flu AD (flu AD = 0)	-2.41	0.0159
BL rate	0.10	0.9187
# elderly	0.17	0.8680

The within group effect of the intervention is interpreted from the values of the z statistic and associated p-value of the prepost variable. For the intervention and control groups, z statistics (p-values) of -17.54 (<0.0001) and -20.21 (<0.0001) respectively indicates that the within group effect is statistically significant for both groups at the alpha = 0.05 level. The significant results for the longitudinal prepost effect is similar between the control and intervention groups as indicated in figure 4-20 and also indicated in the negative values of the z statistics for both groups.

Table 4-42. Secondary GP Office Visit Outcome Model Results (AD = no).

Secondary Outcome Model Results for All Periods (Control Group) (Change in GP Office Visit Rates)		
Effect	Z	p-value
PS	-2.60	0.0093
Period	-19.91	<0.0001
Prepost	-20.21	<0.0001
flu AD (flu AD = 0)	-2.62	0.0089
BL rate	-0.99	0.3212
# elderly	-2.73	0.0064

Rheumatologist and GI Specialist Visit Rates

Model development

The analysis of the secondary outcome, the effect of the OA AD intervention on rheumatologist and GI specialist office visit rates was carried out using the same methods as the primary outcome analysis with the data for rheumatologist and GI specialist office visit rates substituted for the COX-2 utilization data (figure 4-21).

Between group results

The z statistic and the significance level (p-value) of each variable from the secondary specialist office visit outcome model (figure 4-21) are listed in table 4-43.

$$\mathbf{Y = \beta_0 + \beta_1(X_1) + \beta_2(X_2) + \beta_3(X_3) + \beta_4(X_4) + \beta_5(X_5) + \beta_6(X_6)}$$

Where;
Y = change in specialist visit rates (periods 3 to 6 (post-intervention)),
X_1 = physician participation in the intervention (0 = no, 1 = yes),
X_2 = PS (range from 0 to 1),
X_3 = experimental time period (period = 3,4,5,6),
X_4 = physician participation in the influenza AD service (0 = no, 1 = yes),
X_5 = physician baseline specialist visit rate (visits / patient, (period = 2)),
X_6 = number of patients in the GP's practice >65 years old

Figure 4-21. Secondary Outcome Model for Specialist Office Visits

The between group effect of the intervention is interpreted from the z statistics and associated p-value of the AD variable. The z statistic and p-value of 1.44 and 0.1498 respectively indicate that the effect is not significant at the alpha = 0.05 level.

Table 4-43. Secondary Specialist Office Visit Model Results (Periods = 3,4,5,6).

Secondary Outcome Model Results for Post-intervention Periods 3 to 6
(Change in Specialist Office Visit Rates)

Effect	Z	p-value
AD (ad = 0)	1.44	0.1498
PS	-5.98	<0.0001
period	-0.04	0.9700
flu AD (flu AD = 0)	-5.43	<0.0001
BL rate	-23.22	<0.0001
# elderly	-3.01	0.0026

The model (figure 4-21) was used to determine intervention effects on each post intervention time period. Only the period from 181 to 270 days (period five) following the intervention showed significant difference between groups at the alpha = 0.05 level. The z-statistic and p-value associated with the intervention effect are 2.10 and 0.0356 respectively (95% CI (0.0001, 0.0022)). In this case, where the analysis only includes one time period, the interpretation of the coefficient estimate is similar to traditional GLM methods. That is, the coefficient estimate of 0.0012 (AD = no) is interpreted as the non-intervention group having measures of average change rate 0.0012 greater visits/patient/GP than the intervention group.

The model in figure 4-21 was used to determine between group differences in the pre-intervention time periods (period = 1, 2). The z statistic and associated significance level of each variable is listed in table 4-44. The results are interpreted in the same manner as the post-intervention results. The z statistic and p-value for the AD variable are -1.29 and 0.1976 respectively. The p-value indicates that the groups are not significantly different in the pre-intervention periods at the alpha = 0.05 level.

Table 4-44. Secondary Specialist Office Visit Outcome Model Results (Periods = 1,2).

Secondary Outcome Model Results for Post-intervention Periods 1 and 2 (Change in Specialist Office Visit Rates)		
Effect	Z	p-value
AD (ad = 0)	-1.29	0.1976
PS	-4.71	<0.0001
period	-0.90	0.3670
flu AD (flu AD = 0)	-3.86	0.0001
BL rate	-9.21	<0.0001
# elderly	-3.51	0.0004

Table 4-45 depicts the least square means for the two groups (AD = yes and AD = no) for each of the six experimental time periods. The least square mean values are also presented in a graph in figure 4-22.

Table 4-46 depicts the unadjusted means and standard deviations for the two groups (AD = yes and AD = no) for each of the six experimental time periods. The unadjusted mean values are also presented in a graph in figure 4-23.

Secondary Outcome (Specialist Visits) Least Square Means by AD Group						
AD group	**period**					
	1	**2**	**3**	**4**	**5**	**6**
0	0.0005	0	0.0007	0.0005	0.0011	0.0001
1	0.0016	0	-0.0002	0.0003	-0.0001	0.0003

Table 4-45. Least Square Means for Change in Specialist Office Visit Rate by Group (visits/patient).

Within group (longitudinal) results

The within group model was the same as the between group model in figure 4-21 except that the AD group variable is replaced by a prepost variable which measures within group differences between change in specialist office visit rates pre-intervention and post-intervention. The model is run two times; once including only the intervention group and once including only the control group. The z statistic and the associated significance level (p-value) of each variable are listed in table 4-47 for the intervention group and table 4-48 for the control group.

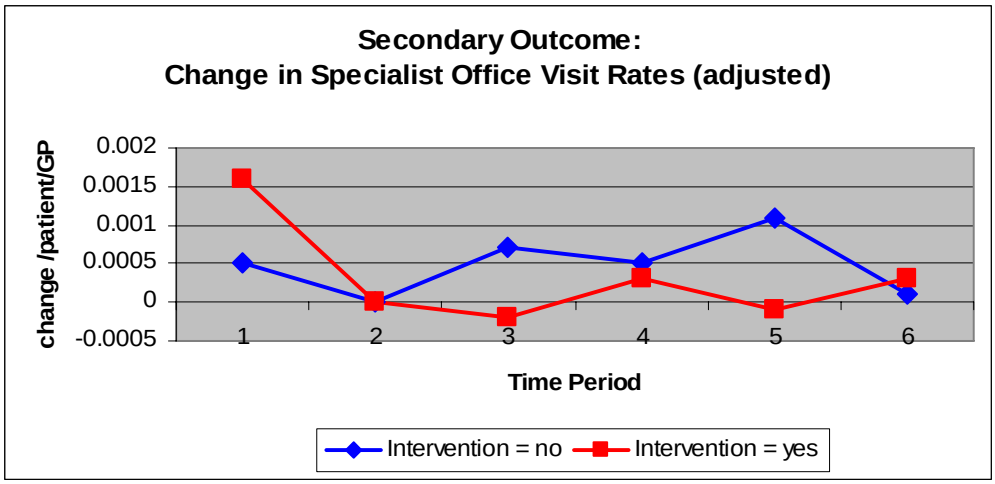

Figure 4-22. Least Square Means for Change in Specialist Office Visit Rates by Group.

Period	AD group			
	0		1	
	Mean	Std Dev	Mean	Std Dev
1	0.0008	0.0089	-0.0002	0.0087
2	0.0000	0.0000	0.0000	0.0000
3	0.0005	0.0077	-0.0011	0.0089
4	0.0008	0.0086	-0.0012	0.0078
5	0.0007	0.0093	-0.0013	0.0079
6	0.0002	0.0083	-0.0006	0.0082

Table 4-46. Unadjusted Means and Standard Deviations for Change in Specialist Office Visit Rate by Group (visits/patient).

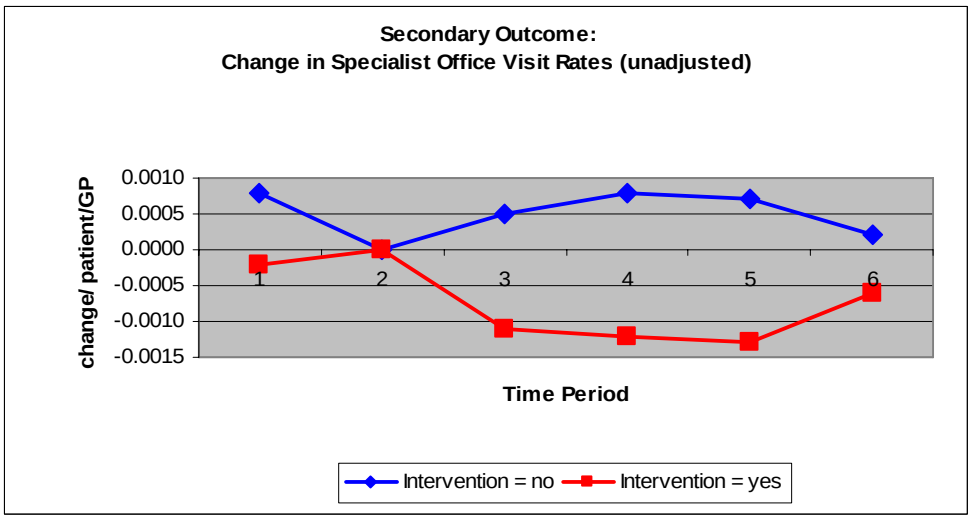

Figure 4-23. Unadjusted Means for Change in Specialist Office Visit Rates by Group.

Table 4-47. Secondary Specialist Office Visit Outcome Model Results (AD = yes).

Secondary Outcome Model Results for All Periods (Intervention Group) (Change in Specialist Office Visit Rates)		
Effect	Z	p-value
PS	-6.45	<0.0001
period	0.87	0.3857
prepost	1.94	0.0519
flu AD (flu AD = 0)	-6.29	<0.0001
BL rate	-17.54	<0.0001
# elderly	-4.56	<0.0001

Table 4-48. Secondary Specialist Office Visit Outcome Model Results (AD = no).

Secondary Outcome Model Results for All Periods (Control Group) (Change in Specialist Office Visit Rates)		
Effect	Z	p-value
PS	-4.24	<0.0001
period	-0.87	0.3870
prepost	-0.70	0.4811
flu AD (flu AD = 0)	-3.25	0.0012
BL rate	-14.67	<0.0001
# elderly	-2.01	0.0444

The within group effect of the intervention is interpreted from the values of the z statistic and associated p-value of the prepost variable. For the intervention and control groups, z statistics (p-values) of 1.94 (0.0519) and –0.70 (0.4811) respectively indicates that the within group effect is not statistically significant for both groups at the alpha = 0.05 level. The results for the longitudinal prepost effect are similar between the control and intervention groups as indicated in figure 4-23.

Hospitalization Rates Due to GI Complications

Model development

The analysis of the secondary outcome, the effect of the OA AD intervention on hospitalization rates was carried out using the same methods as the primary outcome analysis with the data for hospital length of stay rates substituted for the COX-2 utilization data (figure 4-24).

$$Y = \beta_0 + \beta_1(X_1) + \beta_2(X_2) + \beta_3(X_3) + \beta_4(X_4) + \beta_5(X_5) + \beta_6(X_6)$$

Where;
Y = change in hospital utilization rate (periods 3 to 6 (post-intervention)),
X_1 = physician participation in the intervention (0 = no, 1 = yes),
X_2 = PS (range from 0 to 1),
X_3 = experimental time period (period = 3,4,5,6),
X_4 = physician participation in the influenza AD service (0 = no, 1 = yes),
X_5 = physician baseline hospital LOS rate (LOS / patient, (period = 2)),
X_6 = number of patients in the GP's practice >65 years old

Figure 4-24. Secondary Outcome Model for Hospital Length of Stay

Between group results

The z statistic and the significance level (p-value) of each variable from the secondary hospitalization length of stay outcome model (figure 4-24) are listed in table 4-49.

Table 4-49. Secondary Hospital Length of Stay Model Results (Periods = 3,4,5,6).

Secondary Outcome Model Results for Post-intervention Periods 3 to 6 (Change in Hospital Length of Stay)		
Effect	**Z**	**p-value**
AD (AD = 0)	0.33	0.7389
PS	1.48	0.1396
Period	1.15	0.2500
flu AD (flu AD = 0)	1.13	0.2568
Flu AD*quintile (flu AD = 0)	-0.94	0.3468
BL rate	-15.58	<0.0001
los rate	-2.36	0.0183

The between group effect of the intervention is interpreted from the z statistics and associated p-value of the AD variable. The z statistic and p-value of 0.33 and 0.7389 respectively indicate that the effect is not significant at the alpha = 0.05 level.

The model (figure 4-24) was used to determine intervention effects on each post intervention time period. Only the period from 181 to 270 days (period five) following the intervention showed significant difference between groups at the alpha = 0.05 level.

The z-statistic and p-value associated with the intervention effect are 2.49 and 0.0128 respectively (95% CI (1.1093, 3.4627)). In this case, where the analysis only includes one time period, the interpretation of the coefficient estimate is similar to traditional GLM methods. That is, the coefficient estimate of 2.2860 (AD = no) is interpreted as the non-intervention group having measures of average change rate 2.2860 greater visits/patient/GP than the intervention group.

The model in figure 4-18 was used to determine between group differences in the pre-intervention time periods (period = 1, 2). The z statistic and associated significance level of each variable is listed in table 4-50. The results are interpreted in the same manner as the post-intervention results. The z statistic and p-value for the AD variable are 1.58 and 0.1152 respectively. The p-value indicates that the groups are not significantly different in the pre-intervention periods at the alpha = 0.05 level.

Table 4-50. Secondary Hospital Length of Stay Outcome Model Results (Periods = 1,2).

Secondary Outcome Model Results for Post-intervention Periods 1 and 2 (Change in Hospital Length of Stay)

Effect	z	p-value
AD (AD = 0)	1.58	0.1152
quintile	0.56	0.5751
period	-0.67	0.5014
flu AD (flu AD = 0)	-0.10	0.9217
flu AD*quintile (flu AD = 0)	0.72	0.4735
BL rate	-10.32	<0.0001
los rate	-2.84	0.0044

Table 4-51 depicts the least square means for the two groups (AD = yes and AD = no) for each of the six experimental time periods. The least square mean values are also presented in a graph in figure 4-25.

Table 4-52 depicts the unadjusted means and standard deviations for the two groups (AD = yes and AD = no) for each of the six experimental time periods. The unadjusted mean values are also presented in a graph in figure 4-26.

Table 4-51. Least Square Means for Change in Hospital Length of Stay Rates by Group (days/patient).

Secondary Outcome (Hospital LOS) Least Square Means by AD Group						
AD group	period					
	1	2	3	4	5	6
0	0.6168	0	1.0337	0.8165	2.0962	1.7457
1	-0.3396	0	0.2072	1.0211	-0.1898	3.6511

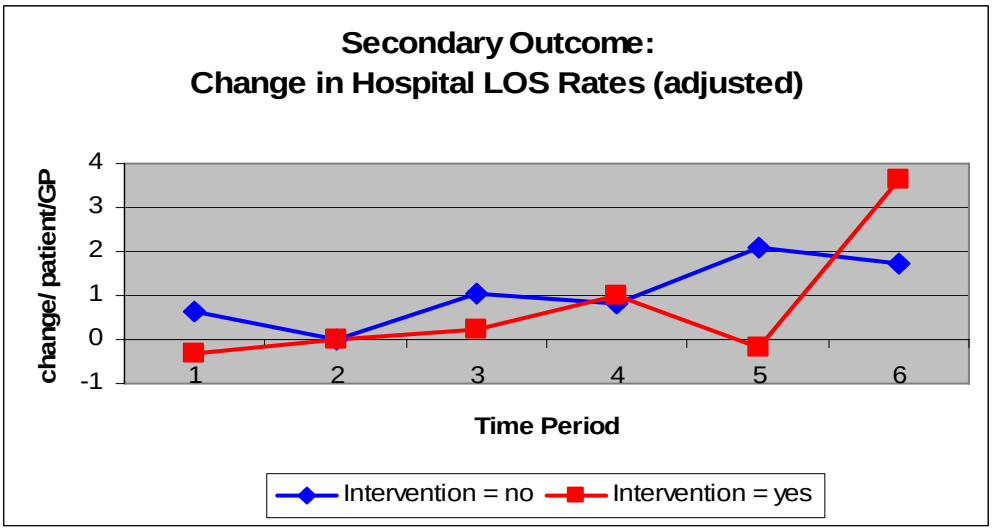

Figure 4-25. Least Square Means for Change in Hospital Length of Stay Rates by Group.

Table 4-52. Unadjusted Means and Standard Deviations for Change in Hospital Length of Stay Rates by Group (days/patient).

Period	AD group			
	0		1	
	Mean	Std Dev	Mean	Std Dev
1	0.7409	8.6598	-0.2482	9.9641
2	0.0000	0.0000	0.0000	0.0000
3	1.3957	9.9291	0.6193	9.7833
4	1.0678	8.4797	1.0061	14.5320
5	1.5773	10.2326	0.5130	8.3747
6	0.5662	8.8979	5.5624	51.6778

Within group (longitudinal) results

The within group model was the same as the between group model in figure 4-24 except that the AD group variable is replaced by a prepost variable which measures within group differences between change in specialist office visit rates pre-intervention and post-intervention. The model is run two times; once including only the intervention group and once including only the control group. The z statistic and the associated significance level (p-value) of each variable are listed in table 4-53 for the intervention group and table 4-54 for the control group.

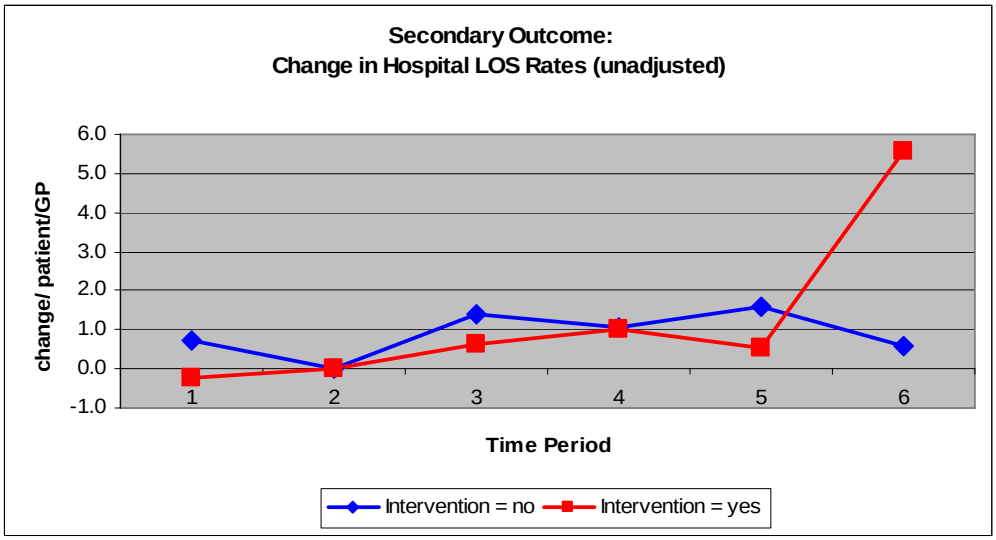

Figure 4-26. Unadjusted Means for Change in Hospital Length of Stay Rates by Group.

Table 4-53. Secondary Hospital Length of Stay Outcome Model Results (AD = yes).

Secondary Outcome Model Results for All Periods (Intervention Group)
(Change in Hospital Length of Stay)

Effect	Z	p-value
PS	0.76	0.4489
period	1.44	0.1491
prepost	0.96	0.3388
flu AD (flu AD = 0)	0.00	0.9976
BL rate	-13.11	<0.0001
# elderly	-1.56	0.1193

Table 4-54. Secondary Hospital Length of Stay Outcome Model Results (AD = no).

Secondary Outcome Model Results for All Periods (Control Group) (Change in Hospital Length of Stay)		
Effect	Z	p-value
PS	3.13	0.0018
period	-1.21	0.2256
prepost	-2.08	0.0375
flu AD (flu AD = 0)	2.99	0.0028
BL rate	-14.40	<0.0001
# elderly	1.87	0.0614

The within group effect of the intervention is interpreted from the values of the z statistic and associated p-value of the prepost variable. For the intervention and control groups, z statistics (p-values) of 0.96 (0.3388) and –2.08 (0.0375) respectively indicates that the within group effect is not statistically significant for the intervention group and is statistically significant for the control group at the alpha = 0.05 level.

Death Rates Due to GI Complications

Model development

The analysis of the secondary outcome, the effect of the OA AD intervention on death rates due to GI complications was carried out using the same methods as the primary outcome analysis with the data for hospital length of stay rates substituted for the COX-2 utilization data (figure 4-27).

$$\mathbf{Ln\ Y = \beta_0 + \beta_1(X_1) + \beta_2(X_2) + \beta_3(X_3) + \beta_4(X_4) + \beta_5(X_5) + \beta_6(X_6)}$$

Where;
Y = death rates (periods 3 to 6 (post-intervention)),
X_1 = physician participation in the intervention (0 = no, 1 = yes),
X_2 = PS (range from 0 to 1),
X_3 = experimental time period (period = 3,4,5,6),
X_4 = physician participation in the influenza AD service (0 = no, 1 = yes),
X_5 = physician baseline hospital LOS rate (LOS / patient, (period = 2)),
X_6 = number of patients in the GP's practice >65 years old

Figure 4-27. Secondary Outcome Model for Deaths Due to GI Complications

Special consideration had to be given to the distribution of the data since the number of deaths per GP per study period was quite low. There were 1984 data points analyzed (496 GPs with six measures each) and in all cases except four the number of deaths per physician was equal to zero or one. The four other cases all contained two deaths (three of the four occurred in the control group). A dichotomous variable representing death/no-death for each period measurement was developed and since the majority of the period measurements represented no-death (142 with death and 2834 without death) a negative binomial distribution was used in the analysis model. The total number of deaths per group per period was less than five in a number of cases. For this reason, the number of study periods was reduced to three by combining periods one and two, three and four, and five and six.

Between and within group results

None of the between or within group analyses of death rates showed significance at the alpha = 0.05 level. The z statistics (p-values) associated with the pre-intervention and post-intervention between group analyses were -0.63 (0.5317) and 0.81 (0.4203) respectively and the z statistics (p-values) associated with the within group analyses for the intervention and control groups were -0.36 (0.7189) and -0.03 (0.9742) respectively.

CHAPTER 5
DISCUSSION

The Academic Detailing Program in Nova Scotia

An analysis of the effect of the OA AD intervention on prescribing behavior should be taken in context of the qualifications of the detailers, the dynamic changes over the course of the intervention and the policy options available to the decision makers. A description of these three topics should add to the determination of generalizability of the intervention to other jurisdictions.

Qualifications of the Detailers

The OA AD intervention employed three detailers; two pharmacists and one registered nurse. One pharmacist worked within the province's capitol district and the other pharmacist and the registered nurse divided the rural area of the province in two. The nurse detailed GPs in the region that she was native to and as such was very familiar with local customs and practices.

All three of the detailers were trained in techniques associated with successful AD programs. These techniques are described in greater detail in appendix A. The intervention was designed to take approximately twenty minutes to present with opportunity for the GP to interact with the detailer over the course of the presentation.

Changes Which Occurred Over the Period of the Intervention (History Effects)

The OA AD intervention was delivered from April, 2002 to November, 2002. The analysis timeframe for our study spanned from October, 2001 (six months before the intervention commenced) to November, 2003 (one year after the intervention concluded).

Between October, 2001 and May, 2003 two warnings regarding the safety of COX-2 inhibitors were issued by Health Canada.[38] The first warning in April, 2002 concerned the results of the VIGOR trial[8] and warned of increased cardiac risk associated with rofecoxib use and the second warning in May, 2002 concerned the results of the CLASS trial[7] and warned of the GI risk associated with celecoxib use particularly in combination with low dose ASA therapy. Analysis of the VIGOR and CLASS trials was included in the OA AD intervention (appendix A). In December, 2004 rofecoxib was withdrawn from the market.[38] The withdrawal occurred after the post intervention study period.

The Nova Scotia pharmacare plan issued a policy change with respect to the benefit status of a combination product containing diclofenac and misoprostol.[39] The product was changed to open benefit status in September, 2002. Announcement of the change was disseminated equally to all GPs in the province. The benefit status of rabeprazole was changed to open benefit in June, 2003[39] after the end of the post intervention analysis period.

Policy Options Available to Decision Makers

Our study examined the effect of the fourth message of the OA AD intervention which addressed pharmacotherapy of OA. The other three messages contained in the intervention were intended to change physician behavior in terms of prescribing non-pharmacologic treatment for OA and research into the effectiveness of these messages is warranted. The OA AD intervention lacked a follow-up visit which is a limitation of the intervention design.[13] Five options available to policy decision makers which could address this shortcoming without the costs associated with a one-on-one follow-up visit are; the distribution of educational material, educational meetings, audit and feedback,

reminders, and changes in benefit schedules.[17, 40] While some of the instruments have not shown significant effects on their own the combination with AD can be effective.[14, 17]

Distribution of educational material

The distribution of educational material involves the dissemination of media (written or video) to the GPs with information reinforcing the messages of the OA AD intervention. It is the decision of the GP to review the message or not. It is relatively low cost and has been shown to have a modest but short-lived effect.[17] The message contained in this medium should be limited to the intervention messages in such a way that does not require "active" learning or interaction with an educator.

Educational meetings

Educational meetings involve meeting in groups to review the messages from the intervention. This instrument can be more complex in nature than the distribution of written material but they are still limited by the inability of the participant to interact with the instructor on a one-to-one basis. Used as a single intervention this instrument has shown little[17] to no effect[14] on improving pharmaceutical use.

Audit and feedback

Audit and feedback is an instrument that involves the analysis of the performance of the provider and/or the provider's peers over a period of time. The instrument is costly to implement as it involves a significant amount of data analysis to produce the audits. Audit and feedback can address some complex issues through the use of the analysis and comparison with peers. Studies using audit and feedback as a single intervention have shown a modest effect.[17]

Reminders and reminder systems

Reminders or reminder systems prompt the provider to recall information. The prompts can take the form of verbal reminders, written notes or computer generated reminders. Reminders of important clinical information in a timely fashion is a benefit to providers however, constant and non-significant alerts (generally from computer systems) create wasted time and can lead to the ignoring of reminders all together. The cost of the chart review for written reminders or the development of clinical software can be quite costly. The effects have been shown to be moderate with statistical significance reached approximately one-half of the time.[17]

Drug benefit changes

Changes to drug benefit schedules can reinforce the intervention's prescribing message by listing some drugs as open benefits where they were restricted before. In terms of the OA AD intervention the listing of a diclofenac and misoprostol combination product and rabeprazole as open benefits[39] coincident with the intervention could encourage prescribing that is in line with the messages contained in the intervention.

Primary Outcome: Effect on COX-2 Utilization Rates

Generalized estimating equations (GEE) techniques for repeated measures were used for the outcomes analyses. The interpretation of the statistical output from the GEE analysis is different from the analysis of output from general linear models (GLM).[37] In GLM with a continuous outcome variable, the coefficient estimate can be interpreted as the effect on the outcome variable if the covariate associated with the coefficient estimate is changed by one unit. In the GEE analysis for repeated measures the main effect result can be interpreted as a between group effect or a within group effect. The magnitude of the contribution of the between and within group effects cannot be determined from the

main effect result alone. To ensure that the between group analysis is indeed showing a significant difference between groups a before and after longitudinal analysis was carried out. The effect of the AD intervention was determined using four separate analyses; a pre-intervention between group analysis, a post-intervention between group analysis, and two within group analyses on the two GP groups.

Statistical Results

The pre-intervention analysis showed that the groups were not significantly different in the six months preceding the intervention ($\underline{z} = 0.88$, $\underline{p} = 0.3775$). The within group analysis for the control group did not show significant change before and after ($z = -0.22$, $p = 0.8273$) whereas the before and after analysis of the intervention group did show a significant decrease in utilization ($z = -2.34$, $p = .0191$). The between group post-intervention analysis by period showed a significant difference between groups on the period immediately following the intervention only ($z = 2.06$, $p = 0.0395$). The intervention was effective in this single period but the between group analysis for the entire post intervention period was not significant ($z = 0.85$, $p = 0.3976$) indicating that the intervention was not sustained beyond the three month post-intervention period.

Practical Significance

The difference in change in COX-2 rates between groups in the period immediately following the intervention is 0.8717 DDDs per patient per 90 days which equals 0.00969 DDDs per patient per day. The inverse of the amount yields the number of patients needed to treat (NNT) to give one patient year of therapy change. The NNT is 104 patients. The average number of elderly patients on a GP panel is 187. The effect can be interpreted as the average GP changing prescribing away from COX-2 inhibitors for 1.8 patients for three months post intervention. This result translates into 416 patients from a

total of 43,197 patients (1% of total) included in the intervention group who had their therapy changed away from COX-2 inhibitors.

Comparison with Literature

The average relative effect of our study on the utilization rate of COX-2 inhibitors over the first 90 day post-intervention period is 23%.

Thomson O'Brien et al. reported that multifaceted AD interventions have shown relative effect sizes ranging from 1 to 45% (from 9 studies).[18] All nine of the studies contained outcomes related to prescribing behavior. In addition to AD the interventions included; provision of educational material (six studies), patient mediated interventions (three studies), social marketing (four studies), audit and feedback (two studies), and reminders (one study). The OA AD intervention employed the provision of educational materials, patient mediated intervention, and a desk top reminder.

Grimshaw et al. reported that AD interventions involving comparisons of process measures showed relative effect sizes ranging from 1.7 to 24% (from six studies) with the median effect equal to 15% and AD interventions involving comparisons of outcome measures showed effect sizes ranging from -1.4 to 13.9% (from 4 studies).[17]

Secondary Outcomes

Effect on Gastro-Protective Agents Utilization Rates

Similar analyses to the primary outcome were carried out on the utilization rates of misoprostol, PPIs, and H2As.

Misoprostol

The misoprostol analyses showed no significant difference between groups in either the pre or post intervention periods (z = -0.87, p = 0.3866 and z = -0.22, p = .8269 respectively) and the longitudinal, within group, analyses (control and intervention)

showed no significant difference within the control or intervention groups over the study period (z = 1.00, p = 0.3176 and z = 0.25, p = 0.8075 respectively). There were no significant between group results by study period.

PPIs

The PPI analyses showed no significant difference between groups in either the pre or post intervention periods (z = 0.13, p = 0.8989 and z = -0.27, p = 0.7906 respectively) and both longitudinal, within group, analyses (control and intervention) showed significant difference (z = -4.22, p = <0.0001 and z = -2.59, p = 0.0097 respectively). The results show a similar increasing utilization pattern for both intervention groups.

H2As

H2As are not indicated as gastro-protective agents and the decision to include this class of medications was made on the basis of the Nova Scotia Pharmacare formulary policy which requires previous authorization for PPIs whereas H2As are an open benefit.

The H2A analyses showed no significant difference between groups in either the pre or post intervention periods (z = 1.09, p = 0.2764 and z = 0.05, p = 0.9619 respectively) and both longitudinal, within group, analyses (control and intervention) showed significant difference (z = -2.33, p = 0.0201 and z = -5.56, p = <0.0001 respectively). The results show a similar increasing utilization pattern for both intervention groups.

Effect on Utilization of Medical Services

Analyses using similar models to the primary outcome were carried out on the change in GP office visit rates, specialist physician visit rates, and hospital length of stay rates. Analysis of the death rates due to GI complications was carried out using similar

statistical methods but the distribution of the data was more consistent with a negative binomial distribution.

The literature suggests the gastro-protective effects of COX-2 inhibitors are removed by daily aspirin therapy[7] and the use of PPIs with traditional NSAIDs provide gastro-protection similar to COX-2 inhibitors.[41] The analysis did not control for these conditions.

GP office visits

The GP office visit analyses showed no significant difference between groups in either the pre or post intervention periods ($z = 0.37$, $p = 0.7097$ and $z = 1.06$, $p = 0.2888$ respectively) and both longitudinal, within group, analyses (control and intervention) showed significant difference ($z = -20.21$, $p = <0.0001$ and $z = -17.54$, $p = <0.0001$ respectively). The results show a similar increasing utilization pattern for both intervention groups. One possible explanation for the increase in GP visits for both groups is a seasonal effect. The time periods with the greatest number of visits coincide with the winter months. The time period from three to six months post intervention showed significantly fewer GP visits in the control group than the intervention group (-0.4195 visits per patient (95% CI (-0.7926, -0.0464)). A possible explanation for the difference is that the intervention GPs monitored patients more closely after the intervention for GI side effects.

Specialist office visits

The specialist office visit analyses showed no significant difference between groups in either the pre or post intervention periods at the alpha = 0.05 level ($z = -1.29$, $p = 0.1976$ and $z = 1.44$, $p = 0.1498$ respectively). The longitudinal, within group, analyses on the control and intervention groups showed no significant difference at the alpha =

0.05 level (z = -0.70, p = 0.4811 and z = 1.94, p = 0.0519 respectively). The time period from six to nine months post intervention showed statistical significance with the control group having higher visit rates (z = 2.10, p = 0.0356, 95% CI (0.0001, 0.0022) visits per patient). While this result is statistically significant the small magnitude of the difference makes it practically insignificant.

Hospitalization rates due to GI side effects

The hospitalization rate due to GI side effects analyses showed no significant difference between groups in either the pre or post intervention periods at the alpha = 0.05 level (z = 1.58, p = 0.1152 and z = 0.33, p = 0.7389 respectively). The longitudinal, within group, analyses on the control group showed significant difference (z = -2.08, p = 0.0375) and the analysis on the intervention group was not significant at the alpha = 0.05 level (z = 0.96, p = 0.3388). The time period from six to nine months post intervention showed statistical significance with the control group having higher hospitalization rates (z = 2.49, p = 0.0128, 95% (CI 1.1093, 3.4627) days per patient). The direction of the effect is opposite to the hypothesis that the intervention group would have higher hospitalization rates. The change in hospitalization rates is intended to indicate the severity of illness due to GI complications and it is not a measure of numbers of patients who experienced adverse GI events. For example, the result showing the control group with higher hospital utilization rates in the period from six to nine months post intervention indicates that more hospital days per elderly patient were attributed to the control group but it does not indicate that more patients in the control group experienced adverse GI events.

Death due to GI complications

None of the between or within group analyses of death rates showed significance at the alpha = 0.05 level. The z statistics (p-values) associated with the pre-intervention and post-intervention between group analyses were -0.63 (0.5317) and 0.81 (0.4203) respectively and the z statistics (p-values) associated with the within group analyses for the intervention and control groups were -0.36 (0.7189) and -0.03 (0.9742) respectively.

Propensity Score Analysis Methods

The pre-PS analysis showed that five of the twelve variables extracted from the administrative data were significantly different at the alpha = 0.05 level. The three PS methods that were carried out as part of this study performed as described in the literature.[11, 12, 16, 27, 28]

"Greedy Matching" Method

The greedy matching method resulted in a 75% reduction in bias on the four VOC. This adjustment for bias was the lowest result of the three methods tested. The method also suffered from a decrease in sample size of 58% which made it unacceptable due to a possible loss in statistical power for the outcomes analysis as well as a loss of generalizability of findings since the tails of the PS distributions would represent the physicians who were eliminated from the study.[12, 28] Parson's uses a case control study example where the controls outnumber the cases 7.4 to 1. Parson's example resulted in 85% of the cases being matched with a control.[28] In our study the ratio of controls to cases is 1.15 to 1 and the percent of matched cases was 45%. The lack of a substantial control group in our study led to the situation where the subjects with the highest and lowest PSs were excluded.

Quintile Method

The stratification by quintile method resulted in an 82% reduction in bias on the four VOC. Rosenbaum and Rubin [11] stated that stratification on quintiles can be expected to remove approximately 90% of the bias for each of the included covariates. D'Agostino[12] included an example where only the covariates with the greatest initial bias were analyzed. The D'Agostino example resulted in an average decrease in bias of 87.4% for the four included variables. This result is similar to our reduction of 82%.

Regression on the PS Method

The regression on the PS method resulted in the greatest reduction in bias on the VOC of 99%. This result combined with the retention of all GP in the model made the regression on the PS the preferred method for our study.

PS Exploratory Analysis

Rosenbaum suggests that the extent to which an unmeasured variable would be balanced by PS methods would be related to the correlation of the unmeasured variable to model covariates.[11] Austin found that the PS method, based on variables extracted from administrative data, was not effective in balancing clinically relevant variables extracted from patient charts but not included in the PS analysis.[29] Our study sought to determine the extent to which a PS model based on administrative data was able to balance administrative variables which were not included in the PS regression model. Our study found that the ability of the PS method to reduce bias on variables not included in the PS regression model was associated with the correlation of the variable which was not included with the PS and included covariates. The reduction of bias on the variable not included in the PS model increased as the correlation between the unmeasured variable and the PS increased. The magnitude of the bias reduction ranged from 39 to 60%

(depending on PS method) when the correlation between PS and the unmeasured variable was 0.22 and the largest correlation with an included covariate was 0.12. The bias reduction increased to 59 to 97% (depending on PS method) when the correlation between PS and the unmeasured variable was 0.33 and the largest correlation with a covariate was 0.91. Our finding supports Rosenbaum's assertion that adjustment for bias on unmeasured variables will increase with the unmeasured variable's correlation with a PS regression model covariate. We also found in one case that the adjustment in bias can be as high as 60% when the correlation with the PS was 0.22 and the highest correlation with a covariate was 0.12.

Limitations

The limitations of our study fall into two general categories; those associated with data and those associated with study design.

Data Limitations

There are a number of database limitations that must be discussed. These limitations can assume the general categories of information that is provided but not 100% reliable and information that is desired but is not captured in the administrative data that is available.

Throughout this research administrative data was relied upon however it is not always accurate and, in fact, its inaccuracy was exploited in one case. ICD-9 codes from hospital discharge data were utilized throughout the secondary outcomes analysis and the reliability of secondary diagnoses, in particular, has been questioned.[42] The events that were evaluated were extracted from the primary, secondary and tertiary diagnoses fields (sixteen diagnosis fields available) and as such are considered to be more reliable.

The PS analysis uses an average hospital length of stay variable as a measure of the wellness of the physician's elderly patient panel. This variable is highly dependent on the individual institutions ability to record admissions and discharges to the institution.

The OA diagnosis variable that was used in the PS analysis takes advantage of the unreliability of ICD-9 coding. Rather than trying to accurately predict the number of patients with OA that a particular physician sees it is used as a measure of the physician's attention to the disease itself.

The aggregated prescription claims data contains inaccuracies due to the fact that it measures the date that a prescription was filled at the pharmacy rather than the date that the prescribing took place and some patients do not have the prescription filled at all. In some cases a considerable time lag may exist between the date the prescription was written and the date it was filled since many patients do not have their prescriptions filled on the day that they were written. The inaccuracies would occur in the instances were the lag time for having a prescription filled caused the data for the claim to be accrued to a study period in which the act of prescribing did not occur and in the instances where prescriptions were not filled.

The prescription claims data supplied by the Nova Scotia Department of Health was subject to a change in encryption methodology carried out by CIHI. These changes lead to the elimination of a data field which indicated whether a prescription was an original fill or a refill. The result was a change in data aggregation for the secondary pharmacotherapeutic outcomes analyses. The refill prescription data was aggregated by the period in which the prescription was dispensed rather than by linkage to the original prescription.

Acetaminophen is not covered under the Nova Scotia pharmacare plan. It has been identified in the OA AD intervention as first line pharmacotherapy for mild to moderate OA[6] but the claims data does not exist so the extent to which this agent is used is not available from the data.

Design Limitations

The omission of a follow-up visit is a weakness in the OA AD intervention. A follow-up visit provides the physician with a boost to their intention to change behavior by presenting the physician with measured actions that reflect the behavioral change that has already taken place.

Grimshaw[17] and Thomson O'Brien[18] identified other non face-to-face follow-up strategies which, if included in a multifaceted intervention, can improve results. Two of the strategies that could be considered as a replacement for the face-to-face follow-up visit are the use of reminder systems (such as chart reminders or computer reminders) to reinforce the original AD messages and physician prescribing profiles (audit and feedback) to give the physician feedback on prescribing performance.

The OA AD intervention does include a reflective exercise that is completed after three months. The reflective exercise requires the physician to re-evaluate the material that was presented in the intervention and it is intended to allow the physician to recommit to his or her intention to change behavior. The exercise also requires the physician to explicitly state actions that will be required to bring about the desired behaviors. The reflective exercise, however, is voluntary and does not involve the academic detailers. There is no assurance that it will be completed and therefore this important component of the intervention may not be realized.

The secondary analyses dealing with utilization of medical services due to GI complications (hospital length of stay, GP and specialist visits) do not include emergency room visits that did not result in a hospital admission.

There are a number of variables that have been cited in other studies as significant indicators for OA pharmacotherapy. Two variables that have been found to be significant are severity of illness of OA and patient pain scale measurements.[10] These variables are not available through the administrative databases for this population and therefore, their omission is a limitation of the study.

At the time of the study, COX-2s were widely used for the treatment of OA. There were other approved uses however (polyps, dysmenhorrea) that could confound the results. This is a limitation of the study however, the effect is expected to be minimal since the other approved uses represent a small percent of the total use and can be expected to be evenly distributed between groups. Off label uses of COX-2s (e.g. rheumatoid arthritis, pancreatic cancer) could also confound results but the effects of these uses are also expected to be minimal.

PS methods themselves can be considered to be a limitation of the study. The theory behind PSs makes the assumption that if the groups are balanced on variables that are measured and relevant then the groups will also be balanced on those variables that are paramount to the study but are not measured.[16] If this theoretical assumption does not hold then the integrity of the quasi-experimental design is in question. This is a limitation of the study and is the rational behind the analysis of the three PS methods to determine if any one method outperforms the others in terms of balancing unmeasured variables.

The PS analyses were based on a review of medical literature and not statistical literature. It is acknowledged that more thoroughly developed PS methodology contained in the statistical literature is not included in our study.

There are threats to the validity of the study from history, maturation and contamination. The history effects that threaten the validity of the study include outside influences on the outcomes such as an increase in GI events due to another cause, new approved uses (labeling changes) from Health Canada during the study period, and new warnings from Health Canada regarding the use of COX-2 inhibitors. The maturation effect exists since the patients in the study are aging over the period of the research. As the patients age their risk for GI events increases and the likelihood of receiving COX-2 therapy also increases. These effects of history and maturation should have the same effect on both of the intervention and control groups and pose a limited threat. The threat from contamination exists since the program is voluntary and there is not a control mechanism in place to prevent physicians for sharing information.

Conclusions

Our study has shown a statistically significant association between an OA AD intervention and the decrease in COX-2 utilization rates in physicians who volunteered for the intervention for the time period immediately following the intervention ($z = 2.06$, $p = 0.0395$, 95% CI (0.0365, 1.4815) DDDs per patient). The positive effect of the intervention remained throughout the post intervention analysis period (figure 4-8) however the between group differences were not statistically significant over the one year post intervention period ($z = 0.85$, $p = 0.3976$). The intervention effect was sustained for three months post intervention.

The relative effect of the between group difference was 23% and the number of patients needed to treat to show a decrease in COX-2 therapy of one DDD is 104.

Since this is an observational study the assertion of a causal relationship is beyond the scope of the research. In an attempt to strengthen the assertion of a causal relationship study design included an intervention and control group and several pre and post intervention time periods were analyzed. The concerns regarding selection bias due to the voluntary nature of the intervention were addressed through the use of regression on PS methods.

The GP office visit between group difference in the time period from three to six months post intervention had practical significance since it showed higher utilization rates in the intervention group ($z = -2.20$, $p = 0.0275$, 95% CI (-0.7926, -0.0464) visits per patient) as compared to the control group. This difference could possibly represent an increased vigilance by the GP towards their patients with respect to GI side effects associated with traditional NSAIDs.

Our study quantified the relationship between the reduction in bias between experimental groups on variables which were not included in the PS regression model and the correlation between the PS and the variable which was not included in the model. The bias reduction on variables not included in the regression analysis, in the selected PS method, was found to range from 60% at a PS correlation of 0.22 and a maximum correlation with an included covariate of 0.12 to 95% at a PS correlation of 0.33 and a maximum correlation with an included covariate of 0.91. This finding is important since it shows that a modest correlation between the variable which was not included in the PS

model and the PS can yield significant reductions in bias between intervention groups on that variable.

Our study was designed and implemented using administrative data to analyze an AD intervention which is part of a continuous program of AD to improve prescribing practices. The methods can be replicated to study the effects of other AD interventions in the same population. Using similar methods for analysis, future research could identify AD topics which have greater or lesser effects on prescribing behavior.

The results from the PS exploratory analysis require further research to generalize our findings in terms of the PS's ability to adjust for bias in unmeasured variables and the magnitude of the adjustment depending on correlation with covariates in the PS regression model.

APPENDIX A
AN APPRAISAL OF THE NOVA SCOTIA OA AD INTERVENTION

In 1990, Soumerai and Avorn summarized eight components that contribute to a successful AD intervention. These components were derived from the literature and from techniques that have been employed and proven by industry for over 100 years.[13] Many of the AD intervention strategies that were identified by Soumerai and Avorn were validated in a review conducted by Davis et al in 1994.[14] The Davis et al literature review analyzed 160 interventions from 99 trials and concluded that the outreach visit, such as AD, is an effective change strategy.[14] Davis et al also concluded that while AD is effective it is seldom used by continuing medical education providers.

The AD intervention on OA that has been developed by the Nova Scotia Department of Health and is managed by the Division of Continuing Medical Education, Dalhousie University, includes all eight components proposed by Soumerai and Avorn.[13]

The eight components are as follows;

- Conduct interviews with physicians to establish baseline knowledge,
- Focus the intervention on specific physicians,
- Define clear objectives for the intervention,
- Establish the credibility of the agency developing the intervention,
- Stimulate physician interaction during the detailing visit,
- Use concise graphic educational materials during the presentation,
- Highlight and reinforce the essential messages during the presentation and,
- Provide positive reinforcement with a follow-up visit to the physician.

The thoroughness of the Nova Scotia OA intervention should be predictive of its success in terms of the work done by Davis et al. The following is a comparative

analysis of the OA AD intervention and the eight components of a successful intervention proposed by Soumerai and Avorn.

Conduct Interviews with Physicians

Before the intervention is developed it is essential to establish baseline knowledge of the targeted physician as well as knowledge of their prescribing behavior and reasons for that behavior.[13]

The establishment of baseline knowledge of the subject will determine at what level the therapeutic teaching portion of the intervention should be targeted. This process will also determine what information the physicians have been given by pharmaceutical company sponsored detailers. The current prescribing behaviors that the physicians exhibit should be analyzed and reasons for these behaviors should be discerned. The collection of this baseline information will contribute greatly toward an intervention that is relevant to the physician group that is being targeted. Since the detailers only have a small amount of time (15 to 30 minutes) with the physician, a direct and poignant information session will be more effective.[13]

Two studies have been included that explicitly describe the process that the authors used to establish the baseline knowledge of their target audience. Ilett employed a pre-intervention survey to determine the needs of the general practitioner population that they were detailing. The study resulted in a significant decrease of 1.4% ($AUS 16,130) in cost of antibiotics within the treatment group.[43] Solomon used physician deviation from guidelines as an indicator that the intervention was warranted. This approach led to a statistically significant 41% decrease in inappropriate prescribing of targeted antibiotics.[44]

The Academic Detailing Service (ADS) has employed a three-tier process in the collection of physician baseline knowledge and behaviors.

- They involve physicians in the selection of subject areas they would like the AD to address.
- They refine this information through teleconferencing with select physicians throughout the province.
- They employ a physician advisory panel to formulate major educational points.

During the solicitation process for participation in an AD session each physician is asked to give feedback on topic areas that they would like to receive AD on in the future. This information is collected and is presented to a group of physicians from throughout the province via teleconferencing. During the influenza vaccine AD intervention that preceded the OA intervention, physicians indicated that they would like to receive information on the management of OA. This information, along with other options was presented to the teleconference physician group and a decision was made to have the physicians' advisory panel develop a list of learning objectives for an OA AD intervention. The Dalhousie CME Division then further developed these major learning points into an academic package for presentation.

Focus Intervention on Specific Physicians

The success of an AD intervention has been attributed to focusing the intervention on certain groups of physicians.[13, 14] In the case of the OA intervention, these groups might include; rheumatologists, physicians with large numbers of elderly patients, or simply physicians whose prescribing patterns differ significantly from the best practice guidelines. These groups should be given extra attention as research has shown that changes in their behavior can have a profound effect on the success of the intervention as a whole.[13]

Five articles described the populations that the interventions were prepared to serve. Ilett and van Eijk targeted the GP population and van Eijk further targeted the pharmacist population.[43, 45] Ilett showed a 1.4% decrease in prescription costs[43] and van Eijk showed a decrease in prescribing of highly anticholinergic antidepressants (not significant) and a significant increase in the prescribing of less anticholinergic antidepressants.[45] Solomon focused the intervention on interns and residents and reported a 41% decrease in inappropriate prescribing of targeted antibiotics.[44] May and Fender targeted their interventions to the individual physician and to the individual physician practice group respectively. May reported a 9% decrease in NSAID prescribing and a decrease in GI events from .20/1000 to .06/1000.[46] Fender reported a statistically significant decrease in specialist referrals (OR = .64) and a significant increase in tranexamic acid prescriptions (OR = 2.38).[47]

The Nova Scotia OA program has targeted the general practitioner population. As a voluntary program the targeting of specific groups within the population would be impractical. The targeting of specific physician groups, such as high variance physicians, is an area for future consideration in the development of new programs.

Define Clear Objectives

The definition of clear objectives is essential in the design of an AD intervention. The objectives should be limited in number (3 or 4) and the outcomes of the objectives should be clearly stipulated and measurement criteria developed. Educational objectives may be to have physicians brought up to date with best practice guidelines but the evidence of the adoption of information may be seen in the measurement of a behavioral objective. Secondary objectives can also be stipulated if they are in line with the overall scope of the intervention.[13]

The definition of clear objectives has been stated in all of the articles that met the inclusion criteria. The objectives were based on clinical guidelines,[43, 44, 48] or primary literature.[45-47]

The Dalhousie AD service on OA has stated four learning objectives.[6] They are;

- Discuss the goals of therapy,
- Recommend non-pharmacological treatments when appropriate,
- Advise patients about the safety and efficacy of acetaminophen, and
- Discuss the role of traditional NSAIDs and COX-2 inhibitors.

The desired behavioral change that is anticipated is the physicians' adherence to best practice guidelines. The behavioral change will be measured through changes in prescribing habits reflected in the provincial and national administrative databases. If guidelines are being followed, then a decrease in prescriptions for both COX-2 inhibitors and NSAIDs is expected. An increase in the use of acetaminophen and non-pharmacological treatments is expected, but unfortunately will not be measured. Indirect measures of appropriate therapy that can be measured through administrative databases include number of visits to primary physicians or specialists and number of hospital visits due to side effects of NSAIDs.

A second outcome of interest, a spin off of optimal therapy, would be a decrease in drug expenditures. In Nova Scotia COX-2 inhibitors are reimbursed under the Seniors Pharmacare Program on a maximum allowable cost (MAC) basis. The MAC is the maximum daily amount that Pharmacare will pay for any drug in that category. Currently, the MAC for COX-2 inhibitors is set at $1.04. Using Celebrex® as an example, if the required daily dose is 400 mg and 100 mg capsules are being supplied the maximum amount per capsule that Pharmacare will pay is $1.04/4 = $0.26. The patient is required to pay the difference in cost between the MAC and the actual cost of the

medication. If a patient therefore, is switched from a COX-2 inhibitor to a traditional NSAID the savings are realized by the patient since the MAC for NSAIDs is set to include full payment for five of the NSAID drugs.

Establish Credibility

There are three criteria that have been identified as being necessary components of a successful AD intervention. The intervention should be produced by an agency that has gained professional respect and whose views are seen to be independent of bias. It should be based on sound evidence from respected sources and academically based educators should present it.[13] Studies that have established credibility by including these criteria have been shown to have statistically significant outcomes.[14]

From the included studies; one stated that the program was developed by a university based expert panel[43] and one stated that the program was developed by the investigator.[45] The programs were presented by a number of different health care professionals including; pharmacists,[43, 44] study team members,[45, 47] clinician educators[44] and physicians.[44, 48]

The AD intervention on OA has been developed through the AD Service of the Continuing Medical Education (CME) Division, Faculty of Medicine, Dalhousie University. Dalhousie CME has a long respected history of providing Maritime physicians with programs designed to improve practice standards.

CME has been offered through the Dalhousie Faculty of Medicine in one form or another for over 50 years. In 1949 the Faculty of Medicine at Dalhousie University began the process of formalizing a program to provide CME to physicians throughout Atlantic Canada. In 1954 funding was obtain from the three Maritime medical societies and in 1957 a division of the Faculty of Medicine was created to administer postgraduate

programs. By 1968 the division had developed a high level of co-operation between small hospitals and was able to deliver innovative and highly relevant programs. The reputation of the Division grew and drew international attention in the form of visits and publications.[6]

In 2001, Dalhousie CME launched the first province wide AD service in Canada. The first topic addressed an update on influenza and pneumococcal vaccines. The service is funded by the Nova Scotia Department of Health and provides physicians in the province with a 15 to 20 minute office visit with a trained health professional on a roughly semi-annual basis.

The second topic addressed the management of OA and it was offered on a voluntary basis over the summer of 2002. The intervention has been developed and is operated by Dalhousie CME, which is under the direction of academically based educators. The intervention's content is entirely evidence-based and the planning committee consists of;

- Two content experts; a local rheumatologist and a drug evaluation pharmacist.
- A family physician advisory panel (three GPs from across the province).
- Three academic detailers; two pharmacists and one registered nurse.

The Dalhousie OA intervention contains many of the criteria that have been identified as contributing to a successful and unbiased detailing service. It is funded by the Nova Scotia Department of Health, an unbiased agency and is operated by the Continuing Medical Education Division of the Faculty of Medicine, Dalhousie University. It is entirely evidence-based and is being presented by academically based health professionals who are given instruction on therapeutic content by the Dalhousie CME Division of the Faculty of Medicine and are specially trained in educational

techniques by the Drug and Therapeutics Information Service (DATIS). DATIS is an internationally recognized organization that provides training in the area of adult education techniques leading to successful medical intervention services.[49]

Stimulate Physician Interaction

Success of an AD intervention has been attributed to the ability of the intervention's presenter to appeal to the physician's own beliefs, needs and values. This can be achieved by engaging the physician in an interactive discussion of the content rather than simply lecturing.[13] The methods used to engage a physician in the process are not often explicitly stated in the literature. There were two cases in the included studies were this component was addressed. In both cases the authors stated that the detailers tailored their presentation to the needs or wants of the physician as a means of stimulating interest.[45, 46] It is, however, a necessary component of a successful intervention.[14]

The Dalhousie OA intervention begins the process of engaging the physician by providing options for additional topics that the detailer can present. The registration page outlines four main messages that will be covered during the visit and it allows the physician to choose any one of seven additional topics that is of particular interest to them. The detailers use this information to tailor a presentation to each physician. The detailers have also been trained in techniques to encourage interactive discussion by DATIS.[49] The flexibility that is built into the OA intervention and the specialized training that the detailers are given should ensure that the physicians are appropriately engaged in the learning experience and that their personal needs are met.

Use Concise Graphic Educational Materials

The use of concise literature as reminders to the physician once the detailing session is completed has been shown to be an effective tool in the success of an intervention. The reminders should be reviewed during the session with the detailer to ensure that the physician receives the proper message.[13] The reminders often take the form of pamphlets, pocket sized aides memoir, graphs, or charts summarizing key points of the presentation. All of the included studies described the physician reminder materials that were developed for their particular intervention.

The Dalhousie OA intervention has accomplished this in two ways. They have produced a lengthy guide to leave with the physician. Highlighted within text boxes in the guide are summary statements to which the physician can easily refer. The detailer also leaves the physician with a laminated 8½ x 11 sheet that summarizes key therapeutic monitoring points of the presentation on one side and provides cost information for different therapies from the perspectives of patients insured under the Nova Scotia Seniors Pharmacare Program and patients with no drug insurance. (Appendix B) For example, drug therapy for a patient who has reached the annual Pharmacare deductible of $350.00 and is prescribed naproxen 500 mg (NSAID) twice daily would cost the government plan $24.46 per month and the patient would pay nothing. If the same patient were prescribed celecoxib 100 mg (COX-2 inhibitor) the charge to the government would be $24.89 per month and the patient would pay $21.78. The message to the physician is clear that if the COX-2 therapy is not indicated then the savings to the patient can be substantial.

The goals of therapy are summarized and reinforced for both the physician and the patient through a desktop pamphlet pad provided by the Nova Scotia Division of The

Arthritis Society. The physician can review the goals of therapy with a patient and tear off a copy for the patient to take home. The OA intervention has partnered with The Arthritis Society to provide a multi-faceted approach to learning and behavior re-enforcement. The combination of clear and concise materials presented to the physician during the OA intervention from both the AD service and The Arthritis Society helps to remind the physician of the intervention's main messages during patient visits.

Highlight and Reinforce Essentials

The repetition of a few major points has been shown to be effective in the presentation of an AD intervention.[13] This is especially true in the practice of medicine where the physician has many different messages presented to him or her relating to many different disease states and therapies. Even if the intervention addresses a very complex issue the main points must be kept to a minimum. If too much is attempted in the short time that the detailer has with the physician, the major points of the intervention and the desired behavioral changes may not be realized. The few primary messages of the intervention should be repeated and summarized throughout the presentation. None of the included studies specifically outlined their methods for reinforcement of the primary messages. One of the articles did state that the central messages were reinforced at the follow-up visit.[47]

The Dalhousie OA intervention has set four primary messages as its objective. These have been discussed earlier under the section - define clear objectives. These four messages are discussed thoroughly in the main text of the document provided to the physician. The goals of therapy have been summarized in a handout format that serves to remind the physician and is available for the patient to take home. The other three primary message of the intervention are summarized in the main document at the end of

the applicable section and are highlighted in a text box for easy reference. The summary messages have also been compiled and placed at the beginning of the main document for quick reference.

Positive Reinforcement with Follow-up

The incorporation of a follow-up visit into the AD plan has been shown to have a two-fold effect on the positive outcome of an intervention.[13] The use of the follow-up visit to reinforce the main messages of the intervention as well as provide positive feedback to the physician has not been universally employed. In cases where a follow-up was planned it was often several months after the initial visit (4 to 6 months) and included feedback to the physician based on data collected since the initial visit.[45-47] In one case a negative event (deviation from hospital established guidelines) triggered a follow-up visit.[44]

The Dalhousie OA intervention has incorporated a face-to-face follow-up visit by the academic detailer but this follow-up visit only takes place if it is requested by the physician. The AD service administrators report that physicians rarely request the follow-up visit. The physician is asked to comment on the usefulness of the intervention using a separate form that is faxed to the Dalhousie CME office. The intervention also offers continuing education credits through the College of Family Physician of Canada if the physician chooses to complete a reflective exercise three months following the detailers visit.

The OA AD intervention is a methodologically sound program. It is evidence based and it meets all but one of the criteria (the provision of a follow-up visit for all participants) that are defined in the literature as being essential components of a successful intervention. The intervention is expected to improve the quality of care given

to patients in Nova Scotia who suffer from OA. It is expected that the cost burden of pharmaco-therapy for the Nova Scotia elderly population will be lessened through the prescribing of equally effective but less expensive agents. It is also expected that the changes in therapy resulting from the intervention will not cause the elderly population any additional morbidity or mortality.

Summary

The OA AD intervention exhibits strength in several areas that will contribute to its success. The Division of CME at Dalhousie University's Faculty of Medicine has ensured that the intervention meets the needs of the province's physicians through teleconferencing with physicians throughout the province and through consultations with a physician advisory panel. The CME division has established itself as a credible source of information for physicians through the provision of educational programs for over 50 years. They have also partnered with the Nova Scotia Division of The Arthritis Society, which is a respected patient advocacy group. The Arthritis Society also adds another facet to the program through their mailings to arthritis patients informing them of the intervention and encouraging them to speak to their physician about their therapy.

The program itself is based on solid clinical evidence obtained from respected peer reviewed journals and therapeutic guidelines. It is designed as an interactive discussion between the physician and the detailer through specialized training given to the detailers by the Australian based DATIS organization. Each physician also has the opportunity to customize the message through an order form that allows the physician to select a number of additional messages that he or she would like the detailer to bring to the session.

The post-intervention components of the intervention include a survey in which the physician reaffirms the desire to modify their behavior to be more in line with the

guidelines that were presented. It also provides the physician with a follow-up reflective exercise to be complete three months after the intervention that again reaffirms the resolve to optimize his or her patients' OA therapy.

The two areas of weakness in the intervention are that it is a voluntary program and it does not include a structured face to face follow-up visit with the participating physicians.

The voluntary nature of the program is static. Physicians in Nova Scotia are free to choose which CME credits they participate in. The perceived weakness would be that the physicians who are already performing at a high level will partake in the intervention and the high variance physicians, who would benefit most from the intervention, will elect to take other forms of CME. This weakness will be addressed in the data analysis, through the use of propensity score methodology. An in-depth description of the propensity score technique is found in the methods section.

The formal follow-up by the detailer is a component that is planned for future interventions but is not included with the OA intervention. This weakness will be especially important in the sustainability of the intervention effect.

APPENDIX B
OA AD DESKTOP REMINDER

Comparative costs of 30 days supply of drugs for OA.

(February 2002)

Drug	Usual daily dose in OA	Approximate prescription price for a person with no drug insurance	Approximate Seniors' Pharmacare co-pay (portion of prescription price paid by senior)	
			Co-pay before the senior reaches the $350/year Pharmacare deductible	Co-pay after the senior reaches the $350/year Pharmacare deductible
Acetaminophen 500mg	2qid	Retail Price ~ 13.77 (including tax) Acetaminophen is not a benefit under most drug plans		
Ibuprofen 400mg	tid	13.13	4.20	0.00
Naproxen 500mg	bid	24.46	7.26	0.00
Ketoprofen 100mg	bid	33.25	9.21	0.00
Flurbiprofen 100mg	bid	30.22	10.03	0.00
Flurbiprofen 50mg	bid	24.55	11.46	4.91
Diclofenac 25mg	tid	26.49	13.20	6.65
Tiaprofenic Acid 300mg	bid	33.70	13.72	3.69
Sulindac 200mg	bid	38.21	18.14	8.11
Diclofenac SR 75mg	bid	43.40	23.32	13.29
Diclofenac 50mg	tid	44.60	24.53	14.50
Etodolac 300mg	bid	45.17	25.09	15.06
Etodolac 200mg	bid	45.77	25.09	15.06
Naproxen 500mg EC	bid	50.53	30.46	20.43
Meloxicam 7.5mg	od	36.08	17.97	9.67
Meloxicam 15mg	od	40.22	16.75	6.80
Celecoxib 100mg	bid	46.67	30.08	21.78
Celecoxib 200mg	od	46.67	30.08	21.78
Rofecoxib 12.5	od	46.67	35.32	29.65
Rofecoxib 25mg	od	46.67	30.08	21.78
Arthrotec 75*	bid	58.26	19.42	0.00
Arthrotec 50*	tid	63.27	21.08	0.00
Misoprostol 200mcg	qid	55.10	16.37	0.00
Omeprazole 20mg**	od	75.17	25.05	0.00

* Arthrotec® is covered under exception status by the Pharmacare programs for the treatment of inflammatory diseases in those patients for whom cytoprotection is required.

** Omeprazole is covered under exception status by the Pharmacare programs.

APPENDIX C
THE THEORETICAL FOUNDATION FOR ACADEMIC DETAILING

Two broad theoretical frameworks that can be used to describe how the academic detailing effect occurs are social theory and theory of planned behavior. The first is social theory[50] which describes how social values of the individual dictate the importance that an individual places on an interaction. The greater the importance attributed to an interaction (social capital) the greater the chance of uptake of the information. The second theory is expected value theory. This theory is illustrated through the use of two behavioral theories: rational decision theory and the theory of planned behavior.

The construct of social theory that will be described here is social capital. The common saying "it's not what you know, but who you know" is an easy way to sum up the construct of social capital as it emphasizes the need for networks to succeed. There are two levels to social capital that are applicable to academic detailing; extra-community networks or "Bridging" and intra-community ties or "Bonding".[51]

In terms of academic detailing, the detailer is the individual that needs to possess social capital in order for the educational visit to be successful and the program itself must be valuable on both the extra-community and intra-community levels.

The social capital that the detailers in Nova Scotia possess on the extra-community level includes their affiliation with the Dalhousie CME Division and their professional credentials (two are pharmacists and one is a registered nurse). The social capital that they possess on the intra-community level involves the development of a network with physicians over time within their assigned region of the province. The program has

extra-community capital due to its evidence based content and community based capital due to the involvement and buy in of general practitioners throughout the province during all stages of the program's development. The program is also developed in conjunction with a local specialist who adds credibility on both levels by adding a local interpretation to the program's evidence base.

The empirical evidence supporting social theory includes studies that explicitly state the credentials of the detailers and the evidence based value of the academic detailing program.[43, 44] Further to this, review articles have summarized that professional and competent detailers, support from trusted institutions, and reliable evidence based information are significant components of a successful academic detailing program.[13, 14]

The methodological challenge for social theory lies in the ability to measure an individual's social capital. In the context of this study the detailers are competent health care professionals who are known to the physicians in their region. An ability to measure one detailer's social capital over another would be valuable in explaining changes that occur in prescribing behavior.

The first of the two expected value theories is a prescriptive or normative theory and is referred to as rational decision theory. It describes how rational decisions are made and the balance between what is desired and what is possible. A particular part of rational decision theory; the expected utility theory will be discussed. The second theory is a descriptive theory that is referred to as psychological decision theory. It goes beyond the rational decision theory and has developed propositions to describe actual behavior.[52]

The proponents of expected utility maintain that it is a normative theory and if the physician adheres to the axioms of the theory each prescribing decision is made by

considering a number of options with expected utility assigned to each. The physician simply chooses the option which maximizes the utility. Two major axioms of utility theory are transitivity and independence. Transitivity states that if A is preferred to B and B is preferred to C then A is preferred to C. Independence states that if A is preferred to B then A with possible consequence C will be preferred to B with the same chance of consequence C.[53]

The critical construct of interest is the transitivity axiom of the theory. The academic detailing program attempts to alter the order of transitivity within the physicians prescribing behavior. If the detailing session is successful in the transfer of the information that COX-2 inhibitors are only as effective for pain relief as traditional NSAIDs or acetaminophen and have limited effect on the reduction of GI events (effective in high risk individuals only) then the cost savings to the patient should place the utility of COX-2 inhibitors lower than that of traditional NSAIDs and acetaminophen. The physician would therefore alter his or her prescribing behavior away from COX-2 inhibitors.

The construct of transitivity within the expected utility theory was chosen because if it holds true it has a direct predictive value on the effect of the academic detailing program on prescribing behavior.

The deviation from the previously described normative behavior is the subject of, and strongest argument for, the theory of planned behavior. The constructs of interest are behavioral intent and perceived behavioral control.[54]

Perceived behavioral control (an individual's perception of their ability to perform a behavior) and intention (an individual's readiness to perform a behavior) are constructs

that have been identified as significant predictors in behavioral change. In order for intent to be manifested into a behavioral change a strong perception of behavioral control must be present.[54]

It is important to note that actual behaviors (behavioral categories) are a collection of single acts and cannot be measured. Single acts include the day to day physician activities of diagnosis, prescribing and referrals and since the collection of single acts make up the behavioral category the measurement of the collection of single acts can act as a proxy for the measurement of the behavioral category. A behavioral act consists of four behavioral elements; an action, target, context, and time.[55]

In the context of this study the four elements in the act of prescribing are an action – the writing of a prescription, a target – the patient, a context – the physician's office, and a time – during a patient visit. The written prescription is a measurable item. The collection of prescriptions is the behavioral category that the academic detailing intends to alter.

Ilett and May have provided examples of studies analyzing the effects of academic detailing that support the proposed theories.[43, 46] Two necessary components of an academic detailing program are that it is evidence based and delivered by credible and trusted detailers. These components support the theories because they put the physician in a position of accepting information that evidence shows should change his or her prescribing behavior. Ilett and May developed interventions based on evidence and the delivery of the program was carried out by reputable agents. Both studies showed significant changes in physician prescribing behavior.[43, 46]

The empirical evidence supporting the theory of planned behavior is illustrated through academic detailing intervention studies that have not been successful in changing physician prescribing behavior.[19, 21] The two studies contained many of the components required for a successful academic detailing intervention however, they were admittedly over ambitious and the message was too complicated for the physician to adopt. The constructs that have been presented would explain the failure of the studies by noting that the intervention may have overwhelmed the physicians and thereby decreased his or her intent to adopt new behavior and the perceived ability to control behavior. If the intent to change and perceived behavioral control are not present then the act of prescribing differently would not be carried out and the effect of the intervention is lost.

The main advantage of the expected utility theory is that it is normative and can be quantitatively measured. The predictive ability of the theory however does not explain why the physician changed his or her behavior. In the context of the osteoarthritis academic detailing intervention the utility theory could be applied to many of the physician patient interactions that lead to the issuing of a prescription. The theory would not explain however deviations from the norm such as prescribing a COX-2 agent to a patient simply because the physician perceives that the patient can afford it. The theory of planned behavior on the other hand does not have the direct predictive power of the utility theory but it explains why the physicians' prescribing behavior is changed. This theory would maintain that a well designed detailing intervention would have the effect of providing the physician with a limited number (3 or 4) of messages thus the physician would perceive his or her ability to change and would intend to change his or her behavior. If this intent is acted upon in a timely fashion then the intent could be

translated into the act of prescribing and the repeated act leads to a change in prescribing behavior.

LIST OF REFERENCES

1. College of Physicians and Surgeons of Nova Scotia. *2002 Annual Listing*. Halifax, NS: College of Physicians and Surgeons of Nova Scotia; 2002.

2. Statistics Canada. 2001 Census of Canada. June 16, 2005; <http://www12.statcan.ca/english/census01/home/index.cfm.> last accessed June 20, 2005.

3. Holbrook AM. *Ontario Treatment Guidelines for Osteoarthritis, Rheumatoid Arthritis, and Acute Musculoskeletal Injury.* Toronto: Ontario Musculoskeletal Therapy Review Panel; 2002.

4. MacLean CH. Quality indicators for the management of osteoarthritis in vulnerable elders. *Ann Intern Med.* Oct 16 2001;135(8 Pt 2):711-721.

5. Lawrence RC, Helmick CG, Arnett FC, et al. Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. *Arthritis Rheum.* May 1998;41(5):778-799.

6. Dalhousie University Continuing Medical Education Website. <http://cme.medicine.dal.ca/ADS.htm> last accessed September 22, 2005.

7. Silverstein FE, Faich G, Goldstein JL, et al. Gastrointestinal toxicity with celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis: the CLASS study: A randomized controlled trial. Celecoxib Long-term Arthritis Safety Study. *JAMA.* Sep 13 2000;284(10):1247-1255.

8. Bombardier C, Laine L, Reicin A, et al. Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis. VIGOR Study Group. *N Engl J Med.* Nov 23 2000;343(21):1520-1528.

9. Cooke C. Utilization of COX-2 Inhibitors in the Nova Scotia Seniors Population: Dalhousie College of Pharmacy; 2001.

10. Cox ER, Motheral B, Frisse M, Behm A, Mager D. Prescribing COX-2s for patients new to cyclo-oxygenase inhibition therapy. *Am J Manag Care.* Nov 2003;9(11):735-742.

11. Rosenbaum PR, Rubin DB. Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association.* September 1984;79(387):516-524.

12. D'Agostino RB, Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med.* Oct 15 1998;17(19):2265-2281.

13. Soumerai SB, Avorn J. Principles of educational outreach ('academic detailing') to improve clinical decision making. *JAMA.* Jan 26 1990;263(4):549-556.

14. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance. A systematic review of the effect of continuing medical education strategies. *JAMA.* Sep 6 1995;274(9):700-705.

15. Seeger JD, Williams PL, Walker AM. An application of propensity score matching using claims data. *Pharmacoepidemiol Drug Saf.* Jan 13 2005.

16. Shadish WR, Cook TD, Campbell DT. *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin; 2001.

17. Grimshaw JM, Thomas RE, MacLennan G, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technol Assess.* Feb 2004;8(6):iii-iv, 1-72.

18. Thomson O'Brien MA, Oxman AD, Davis DA, Haynes RB, Freemantle N, Harvey EL. Educational outreach visits: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev.* 2000(2):CD000409.

19. Lin EH, Simon GE, Katzelnick DJ, Pearson SD. Does physician education on depression management improve treatment in primary care? *J Gen Intern Med.* Sep 2001;16(9):614-619.

20. Lin EH, Katon WJ, Simon GE, et al. Achieving guidelines for the treatment of depression in primary care: is physician education enough? *Med Care.* Aug 1997;35(8):831-842.

21. Brown JB, Shye D, McFarland BH, Nichols GA, Mullooly JP, Johnson RE. Controlled trials of CQI and academic detailing to implement a clinical practice guideline for depression. *Jt Comm J Qual Improv.* Jan 2000;26(1):39-54.

22. Goldberg HI, Wagner EH, Fihn SD, et al. A randomized controlled trial of CQI teams and academic detailing: can they alter compliance with guidelines? *Jt Comm J Qual Improv.* Mar 1998;24(3):130-142.

23. Zwar NA, Wolk J, Gordon JJ, Sanson-Fisher RW. Benzodiazepine prescribing by GP registrars. A trial of educational outreach. *Aust Fam Physician.* Nov 2000;29(11):1104-1107.

24. Tomson Y, Hasselstrom J, Tomson G, Aberg H. Asthma education for Swedish primary care physicians--a study on the effects of "academic detailing" on practice and patient knowledge. *Eur J Clin Pharmacol.* 1997;53(3-4):191-196.

25. Gorins A. Preventing breast cancer. *Eur J Gynaecol Oncol.* 2000;21(3):213.

26. Weitzen S, Lapane KL, Toledano AY, Hume AL, Mor V. Principles for modeling propensity scores in medical research: a systematic literature review. *Pharmacoepidemiol Drug Saf.* Dec 2004;13(12):841-853.

27. Rubin DB. Estimating causal effects from large data sets using propensity scores. *Ann Intern Med.* Oct 15 1997;127(8 Pt 2):757-763.

28. Parsons LS. Reducing Bias in a Propensity Score Matched-Pair Sample Using Greedy Matching Techniques. Paper presented at: 26th Annual SAS Users Group International Conference, 2001; Long Beach, CA.

29. Austin PC, Mamdani MM, Stukel TA, Anderson GM, Tu JV. The use of the propensity score for estimating treatment effects: administrative versus clinical data. *Stat Med.* May 30 2005;24(10):1563-1578.

30. Cepeda MS, Boston R, Farrar JT, Strom BL. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. *Am J Epidemiol.* Aug 1 2003;158(3):280-287.

31. Joffe MM, Rosenbaum PR. Invited commentary: propensity scores. *Am J Epidemiol.* Aug 15 1999;150(4):327-333.

32. Weitzen S, Lapane KL, Toledano AY, Hume AL, Mor V. Weaknesses of goodness-of-fit tests for evaluating propensity score models: the case of the omitted confounder. *Pharmacoepidemiol Drug Saf.* Apr 2005;14(4):227-238.

33. Seeger JD, Walker AM, Williams PL, Saperia GM, Sacks FM. A propensity score-matched cohort study of the effect of statins, mainly fluvastatin, on the occurrence of acute myocardial infarction. *Am J Cardiol.* Dec 15 2003;92(12):1447-1451.

34. Canadian Institute for Health Information. *Privacy and confidentiality of health information at CIHI: principles and policies for the protection of personal health information and policies for institution-identifiable information*. 3rd ed. Ottawa, ON: Canadian Institute for Health Information; 2002.

35. Sas Institute Inc. *SAS/STAT Software: Release 8.2*. Cary, NC: SAS Institute Inc.

36. WHO. Vol 2005: World Health Organisation Collaborating Centre for Drug Statistics Methodology; Oslo, NO; 2005.

37. Twisk JWR. *Applied longitudinal data analysis for epidemiology: a practical guide*. Cambridge, UK; Cambridge University Press; 2003.

38. Health Canada Website. <http://www.hc-sc.gc.ca/ahc-asc/media/advisories-avis/index_e.html> last accessed Dec 02, 2005.

39. Nova Scotia Department of Health Website. <http://www.gov.ns.ca/health/pharmacare> last accessed December 02, 2005.

40. Strom BL. *Pharmacoepidemiology*. 4th ed. Chichester; Hoboken, NJ: J. Wiley; 2005.

41. Wolfe MM, Lichtenstein DR, Singh G. Gastrointestinal toxicity of nonsteroidal antiinflammatory drugs. *N Engl J Med.* Jun 17 1999;340(24):1888-1899.

42. Roos LL, Mustard CA, Nicol JP, et al. Registries and administrative data: organization and accuracy. *Med Care.* Mar 1993;31(3):201-212.

43. Ilett KF, Johnson S, Greenhill G, et al. Modification of general practitioner prescribing of antibiotics by use of a therapeutics adviser (academic detailer). *Br J Clin Pharmacol.* Feb 2000;49(2):168-173.

44. Solomon DH, Van Houten L, Glynn RJ, et al. Academic detailing to improve use of broad-spectrum antibiotics at an academic medical center. *Arch Intern Med.* Aug 13-27 2001;161(15):1897-1902.

45. van Eijk ME, Avorn J, Porsius AJ, de Boer A. Reducing prescribing of highly anticholinergic antidepressants for elderly people: randomised trial of group versus individual academic detailing. *BMJ.* Mar 17 2001;322(7287):654-657.

46. May FW, Rowett DS, Gilbert AL, McNeece JI, Hurley E. Outcomes of an educational-outreach service for community medical practitioners: non-steroidal anti-inflammatory drugs. *Med J Aust.* May 17 1999;170(10):471-474.

47. Fender GR, Prentice A, Gorst T, et al. Randomised controlled trial of educational package on management of menorrhagia in primary care: the Anglia menorrhagia education study. *BMJ.* May 8 1999;318(7193):1246-1250.

48. Denton GD, Smith J, Faust J, Holmboe E. Comparing the efficacy of staff versus housestaff instruction in an intervention to improve hypertension management. *Acad Med.* Dec 2001;76(12):1257-1260.

49. National Prescribing Service Website. <http://www.nps.org.au> last accessed September 22, 2005.

50. Coleman JS. *Foundations of social theory*. Cambridge, MA: Belknap Press of Harvard University Press; 1990.

51. Woolcock M, Narayan D. Social Capital: Implications for Development Theory, Research, and Policy. *World Bank Research Observer.* 2000;15 (2).

52. Kozielecki J. *Psychological decision theory*. Boston, MA: Reidel PWN-Polish Scientific Publishers, 1981.

53. Cohen BJ. Is expected utility theory normative for medical decision making? *Med Decis Making.* Jan-Mar 1996;16(1):1-6.

54. Icek Ajzen Theory of Planned Behavior Website. <http://www-unix.oit.umass.edu/~aizen/tpb.html> last accessed December 05, 2005.

55. Ajzen I. *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall; 1980.

BIOGRAPHICAL SKETCH

Stephen Graham's early education was deeply rooted in the Jesuit tradition. He attended St. Paul's High School and in 1985 he graduated with a Bachelor of Science degree from St. Paul's College at the University of Manitoba, in Winnipeg, Canada.

He received his Air Navigators Wings from the Canadian Forces Air Navigation School and served in the Canadian Air Force as a line navigator until 1992 when began his pharmacy degree at Dalhousie University in Halifax, Nova Scotia, Canada.

Stephen graduated with a Bachelor of Science (pharmacy) degree in 1997 and was employed within the Canadian Forces Medical System until his retirement in 2000.

His academic interests are in the areas of physician behavioral change and in the methodology associated with quasi-experimental design. He plans to contribute to the Canadian health care system through continued work in the areas of health policy and quantitative assessment of medical outcomes.