Methods to Control for Bias in Observational Studies

Anne P. Ehlers

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Public Health

University of Washington

2016

Committee:

David R. Flum

Farhood Farjah

Program Authorized to Offer Degree:

Public Health - Health Services

University of Washington

Abstract

Anne P. Ehlers

Chair of the Supervisory Committee:
Adjunct Professor David R. Flum
Health Services

ABSTRACT

Observational studies often suffer from the problem of confounding, where observed results are biased due to the presence of factors that are strongly associated with both the exposure of interest and the outcome. Typical sources of confounding include factors such as age, sex, and medical comorbidities. The failure to account for confounding in the analytic framework can lead to biased results and ultimately an incorrect inference. Arguably the most common method of accounting for confounding is through the use of regression based approaches, although other methods such as propensity score matching are described.

Beyond confounding, an additional source of bias that must be accounted for is the fact that observational data often is sampled from specified groups of individuals. For example, there may be clusters of individuals who are enrolled in the same health plan or are treated at the same hospital. The effect of this sampling framework is that patient outcomes from one health plan, hospital, etc are correlated. The correlation must be accounted for in the model to account in order to make a correct inference. Models that include multiple levels of analysis (such as patient and hospital) are call multilevel or hierarchical. As with the case of confounding discussed above, there are multiple well described methods to account for unmeasured factors that are contained at the cluster level.

This thesis contains two observational studies that were completed by the author during her course of study in the Master's in Public Health Program. Both studies have been accepted for publication by peer-reviewed journals and this this information is copyrighted.(1,2) These studies will highlight two separate methods to account for confounding, as well as two approaches for hierarchical data analysis.

TABLE OF CONTENTS

STUDY 1: ***Alvimopan Use, Outcomes, and Costs: A Report from the Surgical Care and Outcomes Assessment Program Comparative Effectiveness Research Translation Network Collaborative (1)***

ABSTRACT

Background: Randomized trials show that alvimopan hastens return of bowel function and reduces length of stay by one day among patients undergoing colorectal surgery. However, its effectiveness in routine clinical practice and impact on hospital costs remains uncertain.

Study Design: We performed a retrospective cohort study of patients undergoing elective colorectal surgery in Washington State (2009-2013) using data from a clinical registry (Surgical Care and Outcomes Assessment Program) linked to a statewide hospital discharge database (Comprehensive Hospital Abstract Reporting System). We used generalized estimating equations to evaluate the relationship between alvimopan and outcomes while adjusting for patient, operative, and management characteristics. Hospital charges were converted to costs using hospital-specific charge-to-cost ratios, and were adjusted for inflation to 2013 dollars.

Results: Among 14,781 patients undergoing elective colorectal surgery at 51 hospitals, 1,615 (11%) received alvimopan. Patients who received alvimopan had a LOS that was 1.8 days shorter ($p<0.01$) and costs that were $2,017 lower ($p<0.01$) compared to those who did not receive alvimopan. After adjustment, LOS was 0.9 days shorter ($p<0.01$), and hospital costs were $636 lower ($p=0.02$) among those receiving alvimopan compared to those who did not.

Conclusions: When used in routine clinical practice, alvimopan was associated with a shorter LOS and limited but significant hospital cost savings. Both efficacy and effectiveness data support the use of alvimopan in routine clinical practice, and its use could be measured as a marker of higher quality care.

Introduction

More than 330,000 colorectal operations are performed annually in the United States (US), (3) and approximately 17% of patients will develop a post-operative ileus (POI). (4) POI is estimated to cost the US healthcare system over $1.5 billion per year, (5) which does not include the burden placed on patients and their families. While POI has many potential causes, opioid use is believed to be a key determinant. (6) Alvimopan (Entereg®, Merck) is a pharmaceutical—approved by the Food and Drug Administration (FDA) in 2008—that prevents opioid-induced POI in the setting of colorectal and pelvic surgery. Alvimopan blocks peripheral μ-opioid receptors in the gastrointestinal tract, but has limited systemic absorption and ability to cross the blood brain barrier, and therefore still permits opioid-mediated central pain control. (7) Multiple randomized controlled trials (RCT) have shown that alvimopan decreased time to return of bowel function by 5-28 hours depending on the dose. (8-12) In the one RCT that evaluated LOS, alvimopan reduced LOS by 24 hours. (11)

There remains considerable interest in determining whether alvimopan is effective in routine clinical practice. Many processes that were standardized and carefully monitored in RCTs in order to demonstrate efficacy (such as dose and frequency of drug administration and implementation of enhanced recovery after surgery [ERAS]-type programs) may not occur in routine practice, and thus the drug may appear to be more or less effective than what was initially reported. Concerns over the effectiveness of alvimopan persist as several well-done observational studies have reported LOS savings greater than one day. (13-15) Additionally, some believe that the cost of the drug is excessive and unwarranted even though it is estimated to be approximately $67.50 per pill or $937.50 for a full course of seven days of treatment (16) and few patients actually receive the full course. Both of these factors may prevent adoption of alvimopan in routine clinical care.

To address these questions, we sought to evaluate the relationship between alvimopan use, LOS, and hospital costs among patients undergoing elective colorectal surgery in Washington State (WA). We hypothesized that alvimopan use would be associated with a LOS reduction no greater than one day, and that alvimopan use would not be associated with higher hospital costs.

Methods

We performed a retrospective cohort study of all adult patients undergoing elective colorectal surgery from January 1, 2009 to December 31, 2013 at hospitals participating in the Surgical Care and Outcomes Assessment Program (SCOAP). Patients treated prior to 2009 were excluded as alvimopan was not approved by the FDA until May 2008. SCOAP is a quality improvement and benchmarking collaborative based in WA for which the Comparative Effectiveness Research Translation Network (CERTAIN) provides research and analytic support. (17, 18) This clinical registry collects information about patient demographics, disease characteristics, management, and outcomes. SCOAP data is collected directly from the medical record by trained, audited abstractors using standardized definitions that are available via a secure page at www.SCOAP.org. To obtain hospital cost data, SCOAP cases were linked to the WA Comprehensive Hospital Abstract Reporting System (CHARS), a hospital discharge database that includes data from all public and private hospitals in WA, excluding Veterans Affairs and military hospitals. The use of de-identified data does not require review by the University of Washington Human Subjects Division, and the linkage to CHARS was approved by the WA Department of Social and Health Services Institutional Review Board.

Of patients who had elective colorectal surgery between 2009 and 2013 (n=15,565), the following sequential exclusion criteria were applied: patients who were missing information about alvimopan receipt ($n_{excluded}$=31); patients younger than 18 ($n_{excluded}$ =4); individuals who had colorectal surgery listed as a secondary operation ($n_{excluded}$ =226) or who were designated as having an additional or staged procedure during the same admission ($n_{excluded}$ =67); those who had undergone a colon resection within the previous 30 days ($n_{excluded}$ =70); patient with ASA class V ($n_{excluded}$ =13) or who were intubated ($n_{excluded}$ =3); and those with ASA class listed as "emergent" ($n_{excluded}$ =264) or "not applicable" ($n_{excluded}$ =106) as we could not confirm that these were elective cases.

Potential confounding variables were recorded in SCOAP and included sociodemographic characteristics (age, sex, insurance type, categorized body mass index [BMI]), comorbidities, and management details. A Charlson comorbidity index was calculated for each patient based on comorbidities reported in the medical record at the time of their operation. (19) Other indicators of baseline patient health included the use of home oxygen and mobility devices such as a walker or wheelchair. The primary indication for each operation was classified as diverticulitis, neoplasm (colon cancer, rectal cancer, colon mass, or polyps), inflammatory disease (Crohn's disease or ulcerative colitis), or other. The operation performed was defined by anatomic site (left versus right colon, rectum) and extent of operation (partial versus total colectomy, low anterior resection, abdominoperineal resection, ostomy takedown). Surgical approach was categorized as either open or minimally invasive (including laparoscopic, laparoscopic converted to open, laparoscopic/hand-assisted, robotic or robotic converted to open). Operative time was based on the interval from incision to final wound closure. Patients with operative time less than the 1st percentile, or 40 minutes, were recoded as missing (n=137) as were individuals with operative time greater than 24 hours (n=5). Process-of-care measures included the use of an epidural, patient-controlled analgesia (PCA), and presence a nasogastric tube (NGT) at the conclusion of surgery.

The exposure variable in this study was a binary indicator of alvimopan receipt at any time during the hospital stay. The outcome variables were LOS and total hospital costs. SCOAP records were used to obtain LOS, defined as the total days in hospital from the date of surgery to the date of discharge. No patient was missing LOS data, but 5% of patients (n=760) of patients had a LOS 2 days or shorter and were recoded as missing to prevent the inclusion of potentially erroneous data. Clinical expertise was used to make this judgment. Total charges for the index hospitalization were recorded in CHARS. Payer-perspective costs, adjusted for inflation to 2013 dollars, were derived from hospital charges using publicly available hospital specific Medicaid charge-to-cost ratios. (20, 21) A subset of individuals (n=1,722) was missing charge data if they were treated at a US military hospital (SCOAP collects data from military hospitals even though CHARS does not), a non-WA hospital (SCOAP collects data from one hospital in Oregon as part of a regional quality improvement collaborative), or linkage was not possible.

Patient characteristics were summarized using frequency distributions for categorical variables, and means and medians for continuous variables. Categorical variables were compared using the Pearson chi-square statistic. Continuous variables were compared using the two-tailed Student t-test when normally distributed, and non-parametric equality of medians test when not normally distributed. Multivariable regression using generalized estimating equations (GEE) was used for the primary analysis evaluating the relationship between alvimopan, LOS, and costs after adjusting for potential confounding variables, clustered at the hospital level. We specified a gamma distribution to account for non-normally distributed LOS and costs.(22) A planned, post-hoc analysis included any in-hospital complication as an additional potential confounding variable in the model.

A planned sensitivity analysis was performed used propensity score matching. The probability of receiving alvimopan was estimated using a logistic regression model inclusive of all measured patient and management variables, as well as hospital, listed as potential confounders in the primary analysis. Patients were matched 1:1 using the nearest neighbor method based on propensity score, and LOS and costs were compared between the matched groups, again using GEE to account for correlated data at the hospital level.

Analyses were conducted using STATA, version 12 (STATA Corp, College Station, TX).

Results

From 2009 to 2013, 14,781 patients underwent a colorectal procedure across 51 hospitals. Overall, 1,615 patients (11%) received alvimopan at 26 hospitals that administered the drug at least once. Use among patients increased from 6% in 2009 to 17% in 2013, and the proportion of hospitals where the drug was administered at least once increased from 66% to 82% over the study period. Patients who received alvimopan were more likely to have characteristics typically associated with shorter LOS and lower costs. On average, they were younger; more often had commercial insurance; more frequently had no comorbid conditions and an ASA Class I or II; more frequently had a normal BMI; less frequently used home oxygen or a home mobility device; were less likely to smoke; were more likely to have a neoplasm as an indication for surgery; had higher rates of low anterior resection and laparoscopic surgery; longer median operation time; less frequently had an epidural or NGT at the end of the operation; and more frequently had a PCA. Patients who received alvimopan were significantly less likely to have an in-hospital complication (8% versus 18%, p<0.01). There were no differences with regards to patient sex. (**Table 1**)

Unadjusted analyses showed that patients who received alvimopan had a mean LOS that was 1.8 days shorter (7.0 versus 5.2 days, p<0.01) and mean hospital costs that were $2,017 less ($10,667 versus $12,684, p<0.01) compared to patients who did not receive alvimopan. The differences in patient and management characteristics between patients who did and did not receive the drug were anticipated to have biased the unadjusted outcomes in favor of the alvimopan group. Adjustment for the potential confounding effects of these variables demonstrated a statistically significant association between alvimopan use and shorter LOS of 0.9 days (95% CI -1.1 days, -0.7 days) as well as lower costs (-$636, 95% CI -$1,168, -$105). (**Table 2**)

In the planned post-hoc analysis, which accounted for in-hospital complications, alvimopan receipt remained associated with a significantly shorter LOS (-0.7 days, 95% CI -0.9 days, -0.5 days) but not lower overall hospital costs (-$383, 95% CI -$867, +$101). The planned sensitivity analysis using a propensity score match resulted in a subgroup of 1,668 patients (834 in each cohort), all of whom were treated at a hospital where alvimopan was administered at least once during the study period. After matching, patients who received alvimopan less frequently had an epidural (13% versus 19%, p<0.01), but were otherwise well balanced with regard to patient, operative, and process-of-care characteristics. (**Table 3**) In the propensity score match, LOS was approximately 33 hours shorter (-1.4 days, 95% CI: -1.7 days, -1.0 days) and hospital costs were $1,720 lower (95% CI: -$2,456, -$984) in patients who received alvimopan. **(Table 4)**

Discussion

We sought to describe the effect of alvimopan on LOS and costs among patients undergoing elective colorectal surgery in WA. Knowing that alvimopan decreased LOS by one day in the carefully controlled setting of an RCT, (11) we hypothesized that the LOS savings associated with alvimopan use would be less than one day. In fact, we found this to be the case: patients who received alvimopan had shorter LOS of 0.9 days. Additionally, we found that alvimopan use was associated with limited but significant cost savings of $636. These findings persisted even after several approaches at adjustment for measured potential bias. These findings suggest that, as demonstrated in RCTs, alvimopan use has a modest impact on LOS when used in typical practice. It is important to note the fact that patients who received alvimopan, on average, were lower risk patients compared to those who did not. While the reason for this is unclear, it may be related to use of pre-surgical optimization programs in the pre- that can be implemented along with enhanced recovery protocols.

Previous randomized studies demonstrated that alvimopan leads to reduced incidence of ileus and shorter LOS, (8-12) and observational studies have shown an association between alvimopan use and improved outcomes. (13-15) While these studies consistently showed that alvimopan is associated with either faster return of bowel function or shorter LOS, the magnitude of effect varied between studies. While this variation is expected given differences between RCTs and non-randomized effectiveness evaluations, (23, 24) it complicates the determination of the effectiveness of the drug. Furthermore, the measured outcomes are not consistent across studies (e.g. return of bowel function versus clinical ileus versus LOS) which can make comparisons difficult. We used LOS as a surrogate measure, and previous work has shown that the incidence of POI is associated with longer LOS. (25) Our estimates demonstrated an associated LOS savings of 0.9 days, which is plausible given results from RCTs. (11) While no observational study can definitively prove causality, our findings support the effectiveness of alvimopan in the real-world setting. Despite the fact that these findings were confirmed in the planned sensitivity analysis, the propensity score adjustment resulted in an exaggerated estimate of the association between alvimopan use and LOS. Previous work has demonstrated that propensity score analyses in some cases can inflate rather than mitigate bias due to the presence of unmeasured confounding that cannot be included in the propensity score. (26)

Previous observational studies and modeled analyses of pooled data from RCTs reported cost savings of $731-$1,329 in patients who received alvimopan. (15, 16, 27-29) In our study, use of alvimopan was associated with a cost savings of $636. Due to the fact that charges (and subsequently costs) are not itemized in CHARS, the mechanism of this cost savings is not known. While a $636 dollar cost savings may seem inconsequential for an individual patient, from a healthcare systems perspective it is important. If alvimopan had been given to the 13,166 patients in our study who did not receive it, more than $8 million dollars may have been saved over the five year study period. This is important given the concern that alvimopan might increase the cost of care for patients. A full seven day course of the drug costs nearly $1000 (16) and it must be given to a patient before surgery. Such an intervention is expensive because, like other prophylactic medications given for relatively common events, every patient must be given the drug in order to realize the effects. In the planned sensitivity analysis that included complications, we found that costs were not significantly different between those who did and did not receive alvimopan despite the LOS savings. This may be due to the fact that we included all in-hospital complications, covering a range of events from urinary tract infection to re-operation, introducing significant variability in the associated costs of events. Regardless, alvimopan was not associated with higher costs in this setting further supporting our hypothesis that the drug does not lead to higher costs overall.

Even though we accounted for all measurable sources of bias in this study, we may have overestimated the effect of alvimopan because of an inability to measure all potentially confounding variables. Examples of unmeasured bias may include individual surgeon effects and operative technique (e.g. bowel handling during surgery), which are not captured in clinical databases; optimization of fluids and electrolytes in the peri-operative period; participation in enhanced recovery or fast-track protocols; carbohydrate loading prior to surgery; narcotic minimization strategies; and early resumption of oral intake post-operatively. While omission of these factors may overestimate the impact of alvimopan, there are other factors that may have led us to underestimate the effects of alvimopan. For example, administration of alvimopan was recorded as a binary variable, indicating that the patient received it at least once during their hospital stay. We suspect that in routine clinical practice compliance is lower than what was achieved in the RCTs, something that commonly occurs in the real world setting. (22, 23) The impact of this noncompliance is that patients may not have received the drug as it was designed and therefore the effectiveness of the drug would be less than if it had been correctly administered.

A final consideration is that for patients with LOS of two days or less we recoded them as missing. We made this decision *a priori* based on the empirical data. Anecdotally, it is not unusual for patients in well-established enhanced recovery programs to be discharged on post-operative day two. We found that many of the patients who were discharged on post-operative day two or sooner (meaning within 48 hours of surgery) had undergone extensive procedures such as abdominoperineal resection. Given this finding, we were concerned about the integrity of this data. Post-operative day two was empirically selected as it represented the lowest 5% of the patient population. We did perform a sensitivity analysis to determine the effect of using 24 hours (post-operative day one) as the criteria for "missing data" (as this was the first percentile of data) and found a similar LOS savings associated with alvimopan.

One criticism of alvimopan is that it may simply be a marker for the existence of ERAS programs. While we could not measure the presence or absence of ERAS programs at each hospital, we attempted to account for them by measuring factors that are often included in ERAS pathways such as avoidance of NGT use after surgery and use of directed post-operative pain management strategies intended to avoid ileus.(30) Additionally, a recent observational study of patients undergoing colorectal surgery within the context of an established ERAS program showed that alvimopan use was associated with significantly shorter LOS and lower total costs,(31) supporting the hypothesis that alvimopan is effective outside of the practices associated with ERAS programs. Given this evidence along with the finding that alvimopan has been shown to be associated with lower rates of ileus, (8-12) future studies could explore the association between patient reported outcomes and alvimopan. It remains to be determined if lower rates of ileus would lead to higher patient satisfaction overall. Patient reported outcomes are increasingly identified as an important research target, and such a study may further inform our assessment of the overall value of alvimopan.

Conclusion

This study found that use of alvimopan in WA was associated with shorter LOS and a limited but significant cost savings in patients undergoing elective colorectal surgery, even after accounting for measurable sources of bias. Effectiveness evaluations are important in determining the value of drugs and devices prior to routine adoption and the available data support the use of alvimopan as a quality improvement initiative.

Table 1: Patient, operative and process-of-care variables associated with receipt of alvimopan. Due to rounding, some column percentages may not sum to precisely 100.

	Alvimopan				All		
	No		Yes				p-value
Patient Characteristic	No. N=13,166	%	No. N=1,615	%	No. N=14,781	%	
Median Age [range]	62[18-102]		61[18-96]		62 [18-102]		0.02[1]
Female	7,239	55	869	54	8,108	55	0.37[2]
Private Insurance	6,019	46	856	53	6,875	47	<0.01
Charlson Score[3]							
0	9,260	70	1,216	75	10,476	71	<0.01
1	2,993	23	332	21	3,325	23	
2	731	6	54	3	785	5	
3	182	1	13	1	195	1	
ASA Class							
I	616	5	126	8	742	5	<0.01
II	7,229	55	978	61	8,207	56	
III	4,954	38	492	31	5,446	37	
IV	355	3	17	1	372	3	
BMI[4] 18.5-25	4,201	32	574	36	4,775	32	<0.01
Home Oxygen	179	1	7	<1	186	1	<0.01
Mobility Device[5]	684	5	29	2	703	5	<0.01
Current Smoker	2,665	20	248	15	2,913	20	<0.01
Operative Characteristic							
Laparoscopic[6]	5,663	43	1,030	64	6,693	45	<0.01
Median Operating Time[range]	144[40-795]		150[41-688]		144[40-795]		<0.01
Indication							
Neoplasm	5,854	45	825	51	6,679	45	<0.01
Diverticulitis	3,049	23	379	24	3,428	23	
Inflammatory	1,118	9	133	8	1,251	9	
Other	3,145	24	278	17	3,423	23	
Operation Type[7]							
LAR	4,345	33	729	46	5,074	34	<0.01
Right	4,297	33	364	23	4,661	32	
Left	2,006	15	229	14	2,235	15	
TAC	611	5	88	6	699	5	
APR	549	4	103	6	652	4	
Proctectomy	549	4	37	2	586	4	
Stoma Takedown	715	6	57	4	772	5	
In hospital Complication	2,387	18	131	8	2,518	17	<0.01
Process-of-Care							
Epidural	2,505	19	237	15	2,742	19	<0.01
NG tube[8]	1,192	9	59	4	1,251	9	<0.01
PCA[9]	10,056	76	1,260	78	11,316	77	<0.01

1. *Comparison using test for the median for continuous variables that were not normally distributed.*

2. *Comparison of patients who did and did not receive alvimopan using χ2 tests for heterogeneity unless otherwise indicated.*
3. *Calculated based on reported comorbid conditions and lab values.*
4. *Body mass index (kg/m^2).*
5. *Use of any home mobility device including walker, wheelchair, scooter, or cane.*
6. *Includes laparoscopic, laparoscopic converted to open, laparoscopic/hand-assisted, robotic, and robotic converted to open*
7. *LAR: low anterior resection; Right: right hemicolectomy; Left: left hemicolectomy; TAC: total abdominal colectomy; APR: abdominoperineal resection*
8. *Nasogastric tube in place when the patient left the operating room.*
9. *PCA: patient-controlled analgesia*

Table 2: Unadjusted and adjusted outcomes for length of stay and hospital costs.

	Alvimopan No	Alvimopan Yes	Unadjusted Differences[1]	Adjusted Differences[2]
LOS[3] [95% CI]	7.0 days	5.2 days	-1.8 days [-1.5, -2.1]	-0.9 days[-1.1, -0.7]
Costs[4] [95% CI]	$12,684	$10,667	-$2,017[-$957, -$3,077]	-$636[-$105, -$1,168]

1. *Comparison of patients who did and did not receive alvimopan using two sample t test.*
2. *Adjusted for patient sociodemographic characteristics, clinical comorbidities, operative details, and process-of-care metrics described in* ***Table 1*** *using generalized estimating equation models.*
3. *Length of stay defined as time from surgery to discharge.*
4. *Hospital costs from the payer perspective.*

Table 3: Characteristics of propensity score matched subgroups for planned sensitivity analysis.

	Alvimopan				
	No		Yes		p-value
Patient Characteristic	No. 834	% 50.0	No. 834	% 50.0	
Median Age [range]	60[18,97]		62[18,96]		0.07[1]
Female	467	56	474	57	0.73[2]
Private Insurance	435	52	433	52	0.75
Charlson Score[3]					
0	611	73	618	74	0.74
1	179	22	176	21	
2	34	4	27	3	
3	10	1	13	2	
ASA Class					
1	48	5.8	58	7	0.618
2	507	60.8	489	58.6	
3	271	32.5	276	33.1	
4	8	1	11	1.3	
BMI 18.5-25[4]	291	35	284	34	0.95
Home Oxygen	2	<1	6	1	0.16
Mobility Device[5]	15	2	13	2	0.70
Current Smoker	138	17	142	17	0.79
Operative Characteristic					
Laparoscopic[6]	432	52	397	48	0.09
Median Operating Time[range]	151	[41, 769]	142	[45, 688]	0.04
Indication					
Neoplasm	377	45	360	43	0.16
Diverticulitis	200	23	199	24	
Inflammatory	101	12	85	10	
Other	156	19	190	23	
Operation Extent[7]					
LAR	353	42	334	40	0.17
Right	197	24	214	26	
Left	131	16	121	15	
TAC	65	8	56	7	
APR	61	7	65	8	
Proctectomy	9	1	24	3	
Stoma Takedown	19	2	20	2	
Processes-of-Care					
Epidural	155	19	104	13	<0.01
NG tube[8]	43	5	44	5	0.91
PCA[9]	627	75	644	77	0.33

1. *Comparison using test for the median for continuous variables that were not normally distributed.*
2. *Comparison of patients who did and did not receive alvimopan using χ2 tests for heterogeneity unless otherwise indicated.*
3. *Calculated based on reported comorbid conditions and lab values.*

4. *Body mass index (kg/m^2).*
5. *Use of any home mobility device including walker, wheelchair, scooter, or cane.*
6. *Includes laparoscopic, laparoscopic converted to open, laparoscopic/hand-assisted, robotic, and robotic converted to open*
7. *LAR: low anterior resection; Right: right hemicolectomy; Left: left hemicolectomy; TAC: total abdominal colectomy; APR: abdominoperineal resection*
8. *Nasogastric tube in place when the patient left the operating room.*
9. *PCA: patient-controlled analgesia*

Table 4: Unadjusted and adjusted outcomes for length of stay and hospital costs in the propensity score matched group.

	Alvimopan		**Unadjusted**	**Adjusted**
	No	**Yes**	**Differences[1]**	**Differences[2]**
LOS[3] **[95% CI]**	6.9 days	5.5 days	-1.4 days [-0.9,-1.8]	-1.4 days [-1.7, -1.0]
Costs[4] **[95% CI]**	$12,228	$11,258	-$969 [-$2,253, +$314]	-$1,720 [-$2,456, -$984]

1. *Comparison of patients who did and did not receive alvimopan using two sample t test.*
2. *Adjusted using the generalized estimating equations model to account for correlated data at the hospital level.*
3. *Length of stay defined as time from surgery to discharge.*
4. *Hospital costs from the payer perspective.*

STUDY 1 SUMMARY & TRANSITION

In this first study,(1) the primary analysis employed generalized estimating equations (GEE) to account not only for the multiple sources of measured confounding but also for the fact that outcomes of patients who were treated at the same hospital might be similar to one another in ways that were not measured in the data. (1) In this example, propensity score matching of patients treated at hospitals that administered alvimopan during the study was included as a secondary analysis. Both analyses supported the hypothesis that alvimopan was associated with a shorter LOS but not higher costs.(1)

The next study (2) will again investigate the clinical effectiveness of a drug on patient outcomes. In this case, we studied patients undergoing elective spine surgery and compared those who received intrawound antibiotics (IWA) to those who did not for the purposes of reducing rates of surgical site infection (SSI). As with Study 1,(1) there are multiple sources of measured confounding that are addressed by both logistic regression and propensity score matching, as well as clustered data that is approached under a random-effects framework. This will serve as a comparator of the ways in which these methods can be employed in different study settings.

STUDY 2: ***Intra-Wound Antibiotics and Infection in Spine Fusion Surgery: A Report from Washington State's SCOAP-CERTAIN Collaborative (2)***

ABSTRACT

Background: Surgical site infection (SSI) after spine surgery is classified as a "never event" by the Centers for Medicare and Medicaid. Intra-wound antibiotics (IWA) have been proposed to reduce the incidence of SSI, but robust evidence to support its use is lacking.

Methods: Prospective cohort undergoing spine fusion at twenty Washington State hospitals (July 2011-March 2014) participating in the Spine Surgical Care and Outcomes Assessment Program (Spine SCOAP) linked to a discharge tracking system. Patient, hospital, and operative factors associated with SSI and IWA use during index hospitalizations through 600 days were analyzed using a random-effects logistic model (index), and a time-to-event analysis (follow-up) using Cox Proportional Hazards.

Results: 9,823 patients underwent cervical (47%) or lumbar (53%) procedures (mean age 58, 54% female), with a SSI rate of 1.1% during index hospitalization. Those with SSI were older, more often had diabetes, and more frequently underwent lumbar (versus cervical) fusion compared to those without SSI (all $p<0.01$). Unadjusted rates of SSI during index hospitalization were lower in patients who received IWA (0.8% versus 1.5%). After adjustment for patient, hospital, and operative factors, no benefit was observed in those receiving IWA (OR 0.65, 95% CI: 0.42-1.03). At 12 months, unadjusted rates of SSI were 2.4% and 3.0% for those who did and did not receive antibiotics; after adjustment there was no significant difference (HR 0.94, 95% CI: 0.62-1.42).

Conclusions: While unadjusted analyses indicate a nearly 50% reduction in index SSI using IWA, we did not observe a statistically significant difference after adjustment. Despite its size, this study is underpowered to detect small but potentially relevant improvements in rates of SSI. It remains to be determined if IWA should be promoted as a quality improvement intervention. Concerns related to bias in the use of IWA suggest the benefit of a randomized trial.

Introduction

More than 800,000 spinal operations are performed annually in the United States (US). (32, 33) Surgical site infection (SSI) after spine surgery is a rare but devastating event leading to significant increases in patient morbidity with an estimated cost of $33,000-100,000 per infection. (34, 35) The Center for Medicare and Medicaid Services (CMS) now considers SSI after spine surgery to be a “never event” because it is thought to be largely preventable.(36, 37) Large studies reviewing hospital discharge data show that SSI after spine surgery occurs in approximately 2.5% of patients, (38) but rates as high as 14% have been reported in the literature.(39) In 2009, The National Healthcare Safety Network reported nationally that SSI occurs in 1.0% , 1.5%, and 3.1% of cases after laminectomy, fusion, and revision fusion respectively.(40)

While certain SSI prevention techniques have high levels of evidence and have achieved broad uptake (e.g. antibiotics on time, maintenance of euthermia and euglycemia), other approaches, such as the use of intra-wound antibiotics (IWA), are supported by less compelling data. IWA may alter rates of SSI and are often used in spine surgery patients. The use of IWA involves placing antibiotics, most commonly powdered vancomycin, directly in the surgical wound prior to closure,(41) although there does not appear to be a standardized procedure. It has been shown to be a safe intervention, with few adverse events attributable to IWA use. (34, 41) The first published use of IWA in spine surgery was in 1996, but it likely has been used for several decades prior to this.(42) Only one randomized controlled trial (RCT) has been performed to date, showing no difference in infection rates between patients who did and did not receive IWA (1.61% versus 1.68%, not significant), although the study was underpowered to detect an effect.(43) Meta-analyses of observational studies combined with data from the one RCT suggest that IWA reduces SSI rates, with pooled OR’s of 0.19-0.43 (all $p<0.05$).(15-17) The observation that the effect of IWA on SSI is not observed in the setting of a RCT highlights the possibility of significant confounding and bias in the way IWA is applied. Other issues related to the existing studies of this issue include a lack of information about long term SSI. SSI can occur as late as 12 months or more after skeletal surgery with implants (47, 48) and biofilms that form on implants may contribute to SSI in ways that have not been well assessed. (49) Despite these gaps and the uncertain evidence, clinical guidelines from the North American Spine Society include a recommendation to add IWA "in patients with comorbidities or for those undergoing complicated spine surgery."(50)

We sought to perform an effectiveness evaluation comparing spine surgery patients who did and did not receive IWA in Washington State (WA), monitoring the risk of SSI to see if the effect of IWA changes over time. We hypothesized that the rate of SSI would be lower in patients who received IWA, both during the index hospitalization as well as during the follow-up period

Methods

Many WA hospitals participate in a novel data surveillance and benchmarking network called the Spine Surgical Care and Outcomes Assessment Program (Spine SCOAP), a quality improvement (QI) and benchmarking collaborative.(49) Participation in Spine SCOAP is voluntary, but approximately 75% of eligible spinal procedures in WA take place within the Spine SCOAP network. The Comparative Effectiveness Research Translation Network (CERTAIN) partners with Spine SCOAP to provide research and analytic support once the data have been collected.(18)

All consecutive adult patients undergoing elective cervical or lumbar spinal fusion in Spine SCOAP hospitals between July 2011 and March 2014 were included in this study. Spine SCOAP relies on trained abstractors at each hospital who individually review clinical records to collect data. The data is collected primarily for QI purposes but the protocol used for abstraction is prospectively developed. The clinical records from SCOAP were linked to records in the WA Comprehensive Hospital Abstract Reporting System (CHARS) to obtain follow-up rates of SSI during subsequent hospitalizations. CHARS is an administrative database derived from discharge records from all public and private hospitals in Washington State, excluding Veterans Affairs and US Military hospitals. The University of Washington Human Subjects Division and the WA Institutional Review Board approved this study.

The exposure of interest was IWA receipt during the time of operation. Use of IWA was abstracted from the medical record and was recorded as a yes/no variable. The dose, timing, and type of antibiotic were not recorded. Patients were excluded from this analysis if they were missing IWA status (1% of cohort); no patient was missing SSI data. The primary outcome measured was SSI during the index hospitalization, defined by documentation of one of the following: antibiotics ordered for presumed infection, antibiotics ordered for confirmed infection, or wound reopening/debridement.

Hospital specific IWA usage and SSI rates were compared to assess overall trends and association between the two. Descriptive statistics were used to analyze baseline characteristics between individuals who did and did not receive IWA, as well as those who did and did not develop SSI. Included in this analysis were variables describing patient characteristics (age, sex, race, body mass index (BMI), insurance status); comorbidities (coronary artery disease, hypertension, asthma, diabetes, sleep apnea, osteoporosis, prior spine surgery, history of infection after prior spine surgery); medications (beta blockers, statin, angiotensin-converting enzyme inhibitors/angiotensin receptor blockers, therapeutic anticoagulation, steroids, narcotics); operative characteristics (operative site, surgeon type, length of operation, surgical approach, invasiveness index); SSI prevention metrics (on-time antibiotics, euthermia, glucose monitoring, appropriate use of insulin); and hospital characteristics (hospital type, presence of surgical residency). Operation site was defined as cervical versus lumbar. The invasiveness index was calculated based on 6 possible interventions on each operated vertebra: anterior/posterior decompression, anterior/posterior fusion, and anterior/posterior instrumentation. Each intervention at each level was scored one point and summed .Previous work has shown that increasing levels of invasiveness are associated with risk of SSI in spine surgery patients.(52,53) Surgeon type was coded as neurosurgeon, orthopedic surgeon, or other/missing. Surgical method was defined as “open” versus “minimally invasive.”

Patient characteristics were stratified by both SSI outcomes and receipt of IWA and summarized using frequency distributions for categorical variables; continuous variables were summarized using means, medians and standard deviations. Categorical variables were compared using Pearson chi-square statistic. Continuous variables were compared using the two-tailed Student t-test.

We used multivariate logistic regression with random effects, clustering by hospital site, to estimate the odds of incident SSI for patients who received IWA compared to those who did not. Given the low rate of SSI in the population, we used a stepwise regression method with a priori selection of variables to select the most important variables to include in the random effects model. Any variable with a univariate association probability above 0.2 was not included in our model unless the variable was known to be clinically significant in the development of SSI. Random effects regression models were used in order to account for variance due to unmeasured factors within a hospital that may systematically affect patient outcomes. To the extent that this is true, models using random effects are expected to provide more precise outcome estimates for correlated data.(54)

A propensity score (PS) match was performed as a sensitivity analysis. Patients were assigned a predictive score from zero to one based on their likelihood of receiving IWA. IWA and non-IWA patients were then matched 1:1 with their nearest neighbor using calipers of width 0.03 based on their PS. Patients who were not matched were not included in the propensity score analysis. Standardized differences were examined for balance in the baseline characteristics between the IWA and matched non-IWA patients. A standardized difference of less than 10% was considered to represent good balance between the two groups.(55) Patients who did and did not receive IWA were then compared to one another for to determine rates of SSI, analyzed using two-tailed paired Student t-test.

The secondary outcome was SSI rates during the follow-up period, up to 600 days post-operatively, after which there were few patients with complete follow-up data. In addition to the definitions previously described for SSI at index hospitalization, International Statistical Classification of Diseases and Health Related Problems (ICD-9) codes were used to identify SSI occurrences (**Appendix**). Kaplan-Meier curves, stratified by IWA receipt, were constructed to compare rates of SSI over time to take into account censoring due to differential follow-up. A Cox Proportional Hazard model clustered at the hospital level was used to adjust for differences between those who did and did not receive IWA. The model included patient characteristics (age, sex, race, insurance status, BMI, diagnosis, albumin, and diabetes, American Society of Anesthesia class (ASA class), smoking status), operative characteristics (indication, surgical method, length of operation), and hospital characteristics (type of hospital, presence of surgical residency).

All statistical analyses were performed using Stata version 11.0 (StataCorp, College Station, Texas). All tests were two sided and p-values <0.05 were considered statistically significant.

Results

From July 2011-March 2014, 9,823 patients underwent cervical or lumbar fusion across 20 hospitals. The mean age of subjects was 58 years (SD 12.8), 54% were female. Overall, 55% of patients received IWA, but this varied widely across hospitals (range 10%-98%). Surgical volume varied across hospitals (range 5-1512 cases).(**Figure 1**)

Overall, 111 (1.1%) patients developed SSI. The rate of SSI in patients who did and did not receive IWA was 0.8% and 1.5%, respectively (p<0.01). Rates of SSI varied by hospital (range 0 to 3.2%) but there was no association between IWA use and SSI at the hospital level (p=0.08) or number of procedures performed and SSI rate (p=0.10).

Patients with SSI were older (64 versus 58); had higher BMI (32 versus 30); were less often white (78% versus 87%); more frequently had albumin ≤ 3.5mg/dl (11% versus 3%); more often had diabetes (32% versus 17%) and hypertension (69% versus 52%); more frequently used beta blockers (32% versus 21%); more frequently underwent a lumbar procedure (76% versus 52%) versus a cervical procedure and had operative duration longer than four hours (42% versus 20%); and more frequently were treated at an academic hospital (90% versus 66%) (all p<0.01). There were no differences with regard to sex, insurance status, smoking status, rates of degenerative disc disease, prior surgery, asthma, sleep apnea, coronary artery disease, previous SSI, use of home statins, ACE-I/ARB use, therapeutic anticoagulation, steroid use, or baseline narcotic use. Infection prophylaxis measures such as maintenance of euglycemia, euthermia, and timely administration of prophylactic antibiotics were similar between the two groups, as were operative characteristics such as approach (open versus laparoscopic), invasiveness index, and surgeon type. (**Table 1**)

Patients who received IWA more frequently had private insurance (71% versus 67%); more often were treated by a neurosurgeon (63% versus 44%); had lower rates of degenerative disc disease (21% versus 24%), prior surgery (16% versus 18%), and smoking (19% versus 22%); had lower rates of sleep apnea (14% versus 18%); and had higher rates of baseline narcotic use (57% versus 50%) (all $p<0.01$). Although statistically significant, BMI was similar between the two groups (30.2 versus 29.8, $p<0.01$). There were no significant differences with regards to age, sex, race, albumin status, hypertension, diabetes, asthma, previous surgical site infection, and use of non-narcotic home medications between the two groups. Patients who received IWA less frequently had glucose tested among diabetics (67% versus 76%, $p<0.01$) but were otherwise similar with regards to maintenance of euthermia and on-time antibiotics. There were no significant differences with regards to invasiveness index or anatomic location. (**Table 2**)

After adjustment for patient, operative, and hospital characteristics in the random effects model, IWA was associated with a non-significant 35% reduction in SSI (OR 0.65, 95% CI: 0.42, 1.03). Risk factors for SSI included increasing age (OR 1.03, 95% CI 1.01, 1.05) albumin ≤3.5mg/dl (OR 2.08, 95% CI: 1.04, 4.17), diabetes (OR 1.68, 95% CI: 1.06, 2.67), operation time longer than four hours (OR 2.65, 95% CI 1.23, 5.74), and treatment at an academic hospital (OR 2.63, 95% CI: 1.14, 6.08). Sex, race, insurance status, smoking status, BMI, surgical location, surgical approach, ASA class, presence of a surgical residency program, and invasiveness index score were not associated with increased SSI risk. (**Table 3**).

PS matching found 6,910 patients who did and did not receive IWA with similar characteristics (3,455 in each cohort). After matching, the only significant differences between the two groups were surgeon and hospital type: patients who received IWA were more likely to be treated by an orthopedic surgeon (50% versus 47%, $p=0.02$) or at a hospital with a surgical residency program (30% versus 28%, $p=0.04$). We did not observe a statistically significant difference in rates of SSI between PS matched patients who did and did not receive IWA (0.93% versus 1.30%, $p=0.14$).

The secondary analysis of rates of SSI during the follow-up period included 7,179 patients as we were limited to evaluation of those patients with data linkable to CHARS. At 30 days post-operatively, patients who received IWA had non-significantly lower rates of SSI compared to those who did not (1.8% versus 2.4%, p=0.09). These trends persisted at 90 days (2.2% versus 2.8%, p=0.06), 180 days (2.2% versus 2.9%, p=0.06), and 600 days (2.7% versus 3.6%, p=0.06). After adjustment for patient, operative, and hospital characteristics in the Cox Proportional Hazard Model we did not observe a significant difference between the two groups. The hazard of SSI was 0.94 (95% CI 0.62-1.42) for patients who received IWA compared to those who did not.

Discussion

SSI after spine fusion is a challenging but avoidable complication to patients and the healthcare system. Recent payment policy changes have increased pressure on hospitals to employ risk reducing interventions for SSI, but the evidence base for SSI prevention in spine surgery is limited. In this study of patients within Spine SCOAP, those undergoing cervical or lumbar fusion in WA had an overall SSI rate of 1.1%. While unadjusted rates of SSI were lower in patients who received IWA, those who received IWA were also at lower risk for SSI development prior to surgery. After adjustment for patient, hospital, and operative factors, the odds of SSI and rates of SSI in patients with IWA were not statistically lower than in patients who did not receive IWA. Of particular note, especially given the interest in *hospital-level* QI interventions, we did not find that the increased use of IWA within hospitals was associated with lower rates of SSI. Both groups of patients also demonstrated increasing rates of SSI outside of the 30-day time frame, with approximately 80% of infections occurring by 90 days and the residual by 12 months. This suggests that the 30-day window included in many SSI assessment schemes may be inadequate in assessing SSI following spine surgery.

Although SSI after spine surgery is an infrequent event, it is recognized as a significant problem. SSI can predispose susceptible patients to failure of fusion, sometimes requiring revision instrumentation procedures many years after the index operation.(56) Furthermore, calculating precise rates of SSI in the population at large can be challenging. Using billing claims that are of questionable validity for certain diagnoses,(57) investigators using the Healthcare Cost and Utilization Project found rates of SSI in spine surgery patients to be 2.5% (as defined by a single ICD-9 code of 998.59), leading to significantly higher costs and length of stay.(38) While limited data exists, the cost of SSI following traumatic or elective spine surgery costs between $33,000 and $100,000, respectively.(34,35) If there are 800,000 spine procedures each year,(33) the impact on the healthcare system could be as high as $660 million to $2 billion annually.

Previous research on the effectiveness of IWA has shown a mixed effect with overall benefit among pooled results from observational, uncontrolled trials and one underpowered RCT.(43-46) The additional morbidity associated with SSI, along with the high cost associated with SSI, may explain why the use of IWA is so common (55% within Spine SCOAP). It is also unclear whether surgeons who use IWA incorporate it as standard practice, or whether they select patients at higher risk for IWA use. We found significant differences between IWA and non-IWA patients that would suggest differential use, but the patients receiving IWA were surprisingly more likely to have attributes that placed them at lower risk for SSI, such as higher rates of private insurance, lower rates of smoking, and lower rates of prior surgery. This is contrary to the recommendations from the North American Spine Society (50) and may reflect surgeon or site preference, rather than patient risk stratification. These risk factors need to be accounted for when comparing the rates of SSI, and ideally controlled for in a large randomized trial. The size of such a trial makes it a challenge to fund and organize, but it is necessary because SSI occurs so infrequently. Another complicating issue of a trial of IWA for SSI is that the time window for evaluation of SSI needs to extend beyond the convenient window of 30 days post operatively that is typically tracked. This study found that the effects of IWA do not appear to vary over significantly over time but that complete capture of SSI rates is dependent on having an adequate follow-up period.

This study has several limitations. One consideration is that we did not record the specific antibiotic used, timing of the use during the operation, dose, volume and location of IWA use. A more complete analysis would discuss the variability in dose, volume, and location between surgeons and centers. While 1g of vancomycin powder is the most commonly reported antibiotic and dose reported in the literature, (43-46) it is possible that other antibiotics are used within the Spine SCOAP network which may affect our study outcomes. A second limitation is that the definition of SSI used for this analysis may not accurately reflect the true rates of SSI. In-hospital assessments of SSI in Spine SCOAP follow closely, but are not identical, to those used by the Centers for Disease Control. The definition of SSI that occurred after the index hospitalization may be even more problematic because the sensitivity of ICD-9 diagnostic code schema for SSI in spine patients has not been validated. Some SSIs may not be treated as in-patients and will have been missed by this approach. Some of the hospitalizations with the ICD-9 diagnostic codes used in this schema may reflect other SSIs unrelated to the index surgery. To account for this we did perform a sensitivity analysis using a restricted set of diagnostic codes and found similar rates of SSI. This study is also underpowered to detect the 30% reduction in SSI that we observed which may be clinically significant. In order to accurately detect a 30% decrease in baseline SSI rate (1.1% to 0.77%), we would need a study with more than 36,000 patients. Lastly, the time to event analysis was restricted to a subset of patients with data linkable to CHARS and may introduce bias that is difficult to assess.

Conclusion

In conclusion, this study did not observe significant differences in adjusted rates of SSI between patients who did and did not receive IWA. However, this study is underpowered to detect small but potentially important differences in SSI rates. There were significant differences in the types of patients receiving IWA, the clinicians using IWA and the centers where procedures in which IWA was used are performed. Future studies should be randomized and should characterize the dose, timing, and type of antibiotic use, as well as patient characteristics related to effectiveness of IWA. This is especially the case given the importance of this problem, the perception of benefit of IWA, and the increasing use of this practice for the growing numbers of patients undergoing spine surgery around the world.

Table 1: Patient, operative, and hospital characteristics associated with surgical site infection during index hospitalization

	No Surgical Site Infection N=9,712		Surgical Site Infection N=111		p-value
	Count	%	Count	%	
Patient Characteristics					
Mean Age (years)		58		64	<0.01[1]
Male	4497	46%	44	40%	0.16[2]
Mean Body Mass Index (kg/m^2)		30		32	0.01
Race White	8497	87%	87	78%	<0.01
Private Insurance	6677	69%	72	65%	0.39
Current Smoker	8730	20%	98	15%	0.35
Low albumin (≤3.5mg/dl)	301	3%	12	11%	<0.01
Comorbid Conditions					
Degenerative Disc Disease	2213	23%	21	19%	0.34
Prior Surgery	1643	17%	21	19%	0.58
Hypertension	5080	52%	77	69%	<0.01
Diabetes	1651	17%	36	32%	<0.01
Asthma	1392	14%	20	18%	0.26
Sleep Apnea	1516	16%	21	19%	0.32
Coronary Artery Disease	828	9%	14	13%	0.12
SSI[3] after Previous Spine Surgery	53	1%	2	2%	0.08
Osteoporosis	201	2%	8	8%	<0.01
Home Medications					
Statin	3085	32%	43	39%	0.12
Beta Blocker	2005	21%	36	32%	<0.01
ACE-I[4] or ARB[5]	3252	33%	46	41%	0.08
Therapeutic anticoagulation	253	3%	6	5%	0.07
Steroids	367	4%	2	2%	0.28
Narcotics	5245	54%	57	51%	0.58
Infection Prophylaxis Measures					
Maintenance of Euglycemia					
Insulin started for hyperglycemia[6]	211	45%	4	31%	0.31
Glucose tested among diabetics	1169	71%	29	81%	0.21
On-Time Antibiotics[7]	5125	99%	66	100%	0.52
Euthermia (36+)	8902	97%	104	98%	0.56
Operative/Hospital Characteristics					
Open Surgical Approach	9139	94%	101	91%	0.16
Invasiveness Index[8]		9.13		9.88	0.06
Anatomic Location					
Cervical	4619	48%	27	24%	<0.01
Lumbar	5093	52%	84	76%	

Surgeon Type					
Neurosurgeon	5254	54%	51	46%	0.15
Orthopedic Surgeon	4377	45%	58	52%	
Unknown	81	1%	2	2%	
OR duration (hours) %					
≤2 hours	3403	35%	15	14%	<0.01
2-4 hours	4325	45%	49	44%	
>4 hours	1944	20%	47	42%	
Academic Hospital	6369	66%	100	90%	<0.01

1. *Comparison using Students t-test for continuous variables.*
2. *Comparison of patients who did and did not receive intra-wound antibiotics using χ2 tests for heterogeneity unless otherwise indicated.*
3. *Surgical site infection.*
4. *Angiotensin converting enzyme inhibitor.*
5. *Antiogensin II receptor blockers.*
6. *Hyperglycemia defined as >180mg/dl*
7. *This number describes the proportion of patients who received on-time antibiotics, among those who had this recorded. This was not included as a metric in 2013.*
8. *Based on number of instrumentation levels, decompressive levels and fusion levels.*

Table 2: Patient, operative, and hospital characteristics associated with receipt of intra-wound antibiotics during their initial operation.

	No Intra-wound Antibiotics N=4,464		Intra-wound Antibiotics N=5,359		p-value
	Count	%	Count	%	
Patient Characteristics					
Mean Age		58		58	0.80[1]
Male	2084	47%	2457	46%	0.41[2]
Mean Body Mass Index (kg/m^2)		29.8		30.2	<0.01
Race White	3870	87%	4714	88%	0.06
Private Insurance	2970	67%	3779	71%	<0.01
Current Smoker	3995	22%	4833	19%	<0.01
Low albumin (≤3.5mg/dl)	133	3%	180	3%	0.29
Comorbid Conditions					
Degenerative Disc Disease	1093	24%	1141	21%	<0.01
Prior Surgery	807	18%	857	16%	<0.01
Hypertension	2365	53%	2792	52%	0.38
Diabetes	778	17%	909	17%	0.55
Asthma	653	15%	759	14%	0.51
Sleep Apnea	789	18%	748	14%	<0.01
Coronary Artery Disease	412	9%	430	8%	0.04
SSI[3] after Previous Spine Surgery	29	1%	26	0%	0.28
Osteoporosis	111	3%	98	2%	<0.01
Home medications					
Statin	1412	32%	1716	32%	0.68
Beta Blocker	942	21%	1099	21%	0.46
ACE-I[4] or ARB[5]	1479	33%	1819	34%	0.41
Therapeutic anticoagulation	119	3%	140	3%	0.87
Steroids	178	4%	191	4%	0.28
Narcotics	2246	50%	3056	57%	<0.01
Infection Prophylaxis Measures					
Maintenance of Euglycemia					
Insulin started for hyperglycemia[6]	110	45%	105	44%	0.93
Glucose tested among diabetics	588	76%	610	67%	<0.01
On-Time Antibiotics[7]	2289	99%	2902	99%	0.38
Euthermia (36+)	4188	97%	4818	97%	0.73
Operative/Hospital Characteristics					
Open Surgical Approach	4171	94%	5069	95%	0.03
Invasiveness Index[8]		9.11		9.16	0.54
Anatomic Location					
Cervical	2146	48%	2500	47%	0.16

Lumbar	2318	52%	2859	53%	
Surgeon Type					
Neurosurgeon	1940	44%	3365	63%	<0.01
Orthopedic Surgeon	2496	56%	1939	36%	
Unknown	28	1%	55	1%	
Operation duration					
≤2 hours	1449	33%	1969	37%	<0.01
2-4 hours	1929	43%	2445	46%	
>4 hours	1070	24%	921	17%	
Academic Hospital	3224	72%	3245	61%	<0.01

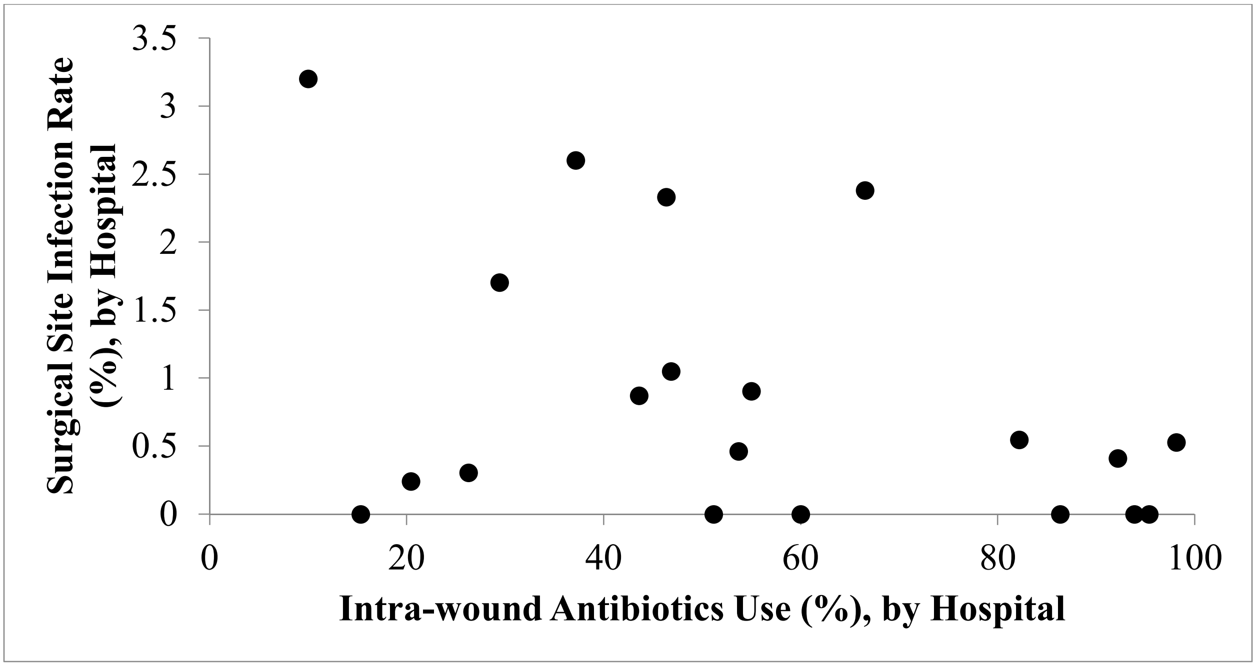
1. *Comparison using Students t-test for continuous variables.*
2. *Comparison of patients who did and did not receive intra-wound antibiotics using χ2 tests for heterogeneity unless otherwise indicated.*
3. *Surgical site infection.*
4. *Angiotensin converting enzyme inhibitor.*
5. *Angiotensin II receptor blockers.*
6. *Hyperglycemia defined as >180mg/dl*
7. *This number describes the proportion of patients who received on-time antibiotics, among those who had this recorded. This was not included as a metric in 2013.*
8. *Based on number of instrumentation levels, decompressive levels and fusion levels.*

Table 3: Odds ratios (index hospitalization) and hazard ratio (follow-up period) for incidence of surgical site infection.

Variable	Unadjusted Odds Ratio	95% CI	Adjusted Odds Ratio	95% CI	Hazard Ratio	95% CI
Intra-Wound Antibiotics	0.67	0.44-1.03	0.65	0.42-1.03	0.94	0.62-1.42
Age	1.04	1.02-1.05	1.03	1.01-1.05	1.02	1.01-1.04
Male	0.74	0.51-1.08	0.68	0.44-1.04	0.78	0.59-1.04
Race White	0.68	0.42-1.08	0.65	0.39-1.10	0.82	0.52-1.04
Current Smoker	0.77	0.45-1.30	1.08	0.60-1.94	1.42	1.13-1.80
Private Insurance	0.85	0.57-1.27	0.87	0.56-1.36	0.70	0.52-0.93
BMI>30[1]	1.45	0.99-2.11	1.35	0.87-2.10	1.31	1.02-1.69
Albumin <3.0	3.62	1.98-6.60	2.08	1.04-4.17	2.39	1.23-4.64
Diabetes	2.32	1.56-3.45	1.68	1.06-2.67	1.15	0.80-1.65
Prior Surgery	1.08	0.67-1.75	0.82	0.48-1.40	0.81	0.55-1.19
Cervical Site[2]	0.41	0.26-0.63	0.68	0.41-1.13	0.54	0.39-0.75
Operating time (hours)						
<2	Ref	Ref	Ref	Ref	Ref	Ref
2-4	1.96	1.13-3.40	1.94	0.98-3.83	1.67	1.07-2.59
>4	3.83	2.14-6.86	2.65	1.23-5.74	2.12	1.51-2.97
ASA[3] Class						
1	Ref	Ref	Ref	Ref	Ref	Ref
2	1.99	0.48-8.27	1.91	0.26-14.28	4.37	0.82-23.20
3	4.84	1.17-19.93	3.05	0.40-23.19	6.39	0.97-42.01
4	15.96	3.11-81.78	7.67	0.80-73.78	10.59	1.47-76.52
Academic Hospital	5.11	1.84-14.20	2.63	1.14-6.08	2.25	1.45-3.49

1. Body mass index, kg/m^2.
2. Versus lumbar site.
3. American Society of Anesthesiologists Class.

Figure 1: Rates of IWA use and surgical site infection across 20 Spine SCOAP Hospitals. No significant association was found between rates of IWA use at the hospital level and rates of SSI (p=0.08).



SUMMARY & CONCLUSION

The studies contained in this thesis (1, 2) provide two examples of how confounding and clustered data can be approached in observational research. Here I will briefly describe the methods that were used in these studies.

Regression models are used when there is reason to believe that the observed treatment effect or outcome is not directly related to the treatment itself; rather, there are other factors such as age, sex, and socioeconomic status that are causing the observed outcome. Regression models seek to approximate the value of some population mean based on a set of known covariates. Under this framework, each predictor is assigned a coefficient that can then be used to construct an equation describing the distribution of the population mean. In the examples in this thesis, we were trying to estimate the effect of a drug on a particular outcome taking into account the other know parameters. However, in both cases we were concerned that there were unmeasured factors that influenced both the treatment assignment and the observed outcome. One of the key points here is that regression analyses are meant to approximate a given *outcome* given a patient's characteristics.

Rather than focusing solely on outcomes, propensity scores approximate an individual's likelihood of receiving the treatment to achieve balance between treated and untreated individuals.(58) Propensity score analyses are often used when there is concern for selection bias in treatment assignment.(59) The goal is to match patients with the same likelihood of receiving treatment to more precisely measure the treatment effect. Under this framework, a logistic regression model predictive of treatment receipt is used to generate a score from zero to one based on the likelihood that a patient would have been treated given their characteristics. The logistic model includes both confounding variables and variables that are associated with the outcome, as these variables may also affect treatment assignment. (58,59) Treated and untreated patients are then matched based on their predicted score to create two cohorts with balanced covariates.(59) Much like regression models, propensity scores can only balance measured sourced of confounding. Propensity scores assume that by balancing measured covariates, you are also able to balance unmeasured covariates. The concern with this assertion is that if propensity scores perfectly predicted treatment assignment, we would not see differences in treatment assignment among two patients with the predicted likelihood of treatment.(60)

Modelled analyses have shown that use of propensity scores may actually increase study bias, rather than mitigate it. Brooks, et al, argue that in order for propensity score matching to achieve its goal, we have to assume that the distribution of any unmeasured covariates is independent of the distribution of measured covariates. Without this, patients with the same propensity score would always have the same treatment assignment. In these modelled analyses, forcing covariate balance among treated and untreated patients results in a greater degree of imbalance in the unmeasured covariates. The effect of this is to increase study bias because the propensity score matched patients are fundamentally different from one another in ways that we cannot measure, and this is not ultimately reflected in the study outcome.(60) In this case, it may be that the resulting imbalance in unmeasured covariates is leading to the difference in outcomes, and not the treatment itself.

In addition to the issue of confounding, many observational datasets contained “clustered” data, where multiple observations are drawn from the same individual, the same neighborhood, etc. In the case of both studies in this thesis, multiple observations (i.e. patient observations) were drawn from the same hospital. The concern with clustered data is that observations from the same cluster tend to be correlated, which goes against traditional assumptions of regression that all observations are independent.(60) Failure to account for this can again lead to a biased estimation and incorrect inference.

Study 1 (1) addressed the clustered data by employing a framework called generalized estimating equations (GEE) which seeks to provide a population average, or marginal interpretation, based on changes in the specified covariates. The variance is then estimated by specifying a working correlation structure based on one’s understanding of the data. If the underlying correlation structure is misspecified, the estimate of the population mean is correct but the standard errors (and thus inference) are not; however, this can be corrected through the use of robust standard errors. The general interpretation of the marginal is the change in the population mean given a change in the covariates.(62)

Study 2 (2) accounted for the clustering by use of a random effects model (also referred to as a mixed model). Under this framework, we assume that each individual cluster, in this case the hospital, is drawn from a larger sample of hospitals whose outcomes are unknown. The model accounts for this by measuring the effect of the intervention within each hospital as well as the effect of the intervention across all hospitals. What this does is calculate cluster-specific means and then uses those means to generate an overall, population-level mean and variance. The resulting model gives an estimate of the treatment effect for the entire study population to provide a more generalizable result. One criticism of the random effects model is that misspecification of the model can lead to incorrect inferences, which GEE is robust to as discussed previously.(61,62)

This thesis provides a broad overview of the types of methods that can be used in the analysis of observational data. There are certainly other methods that are outside of the scope of the studies discussed here which may be more appropriate for the research question. Researchers considering analysis of observational data should consider not only the source of the data, but also potential factors which might bias their results. These should be identified *a priori*, prior to beginning data analysis, along with a proposed analytic plan to promote the integrity of the work.

Acknowledgement

Dr. Ehlers was supported by a training grant from the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health under Award Number T32DK070555. Dr. Farjah received support as a Cancer Research Network Scholar (CRN4: Cancer Research Resources & Collaboration in Integrated Health Care Systems, grant number U24 CA171524). Research reported in this publication was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under Award Number R21AR068632. SCOAP is a program of the Foundation for HealthCare Quality. The Comparative Effectiveness Research Translation Network (CERTAIN) is a program of the University of Washington that provided research and analytic support for this publication and was supported by funding from the Agency for Healthcare Research and Quality under award number R01HS020025. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the Cancer Research Network, the Foundation for HealthCare Quality, or the Agency for Healthcare Research and Quality.

References

1. Ehlers AP, Simianu VV, Bastawrous AL, et al. Alvimopan Use, Outcomes, and Costs: A Report from the SCOAP-CERTAIN Collaborative. Journal of the American College of Surgeons. *In Press.*
2. Ehlers AP, Khor S, Shonnard N, et al. Intra-wound Antibiotics and Infection in Spine Fusion Surgery: A Report from Washington State's SCOAP-CERTAIN Collaborative. Surg Infect (Larchmt). 2016 Apr;17(2):179-86.
3. Weiss AJ, Elixhauser A, Andrews RM. Characteristics of Operating Room Procedures in U.S. Hospitals, 2011: Statistical Brief #170. In: Healthcare Cost and Utilization Project (HCUP) Statistical Briefs. Rockville (MD): ; 2006.
4. Iyer S, Saunders WB, Stemkowski S. Economic burden of postoperative ileus associated with colectomy in the United States. J Manag Care Pharm. 2009 Jul-Aug;15(6):485-94.
5. Goldstein JL, Matuszewski KA, Delaney CP, et al. Economic Burden of Postoperative Ileus Associated with Abdominal Surgery in the United States. P&T. 2007;32(2):82,82-90.
6. Holte K, Kehlet H. Postoperative ileus: a preventable event. Br J Surg. 2000 Nov;87(11):1480-93.
7. Schmidt WK. Alvimopan* (ADL 8-2698) is a novel peripheral opioid antagonist. Am J Surg. 2001 Nov;182(5A Suppl):27S-38S.
8. Wolff BG, Michelassi F, Gerkin TM, et al. Alvimopan, a novel, peripherally acting mu opioid antagonist: results of a multicenter, randomized, double-blind, placebo-controlled, phase III trial of major abdominal surgery and postoperative ileus. Ann Surg. 2004 Oct;240(4):728,34; discussion 734-5.
9. Delaney CP, Weese JL, Hyman NH, et al. Phase III trial of alvimopan, a novel, peripherally acting, mu opioid antagonist, for postoperative ileus after major abdominal surgery. Dis Colon Rectum. 2005 Jun;48(6):1114,25; discussion 1125-6; author reply 1127-9.
10. Viscusi ER, Goldstein S, Witkowski T,et al. Alvimopan, a peripherally acting mu-opioid receptor antagonist, compared with placebo in postoperative ileus after major abdominal surgery: results of a randomized, double-blind, controlled study. Surg Endosc. 2006 Jan;20(1):64-70.
11. Ludwig K, Enker WE, Delaney CP, et al. Gastrointestinal tract recovery in patients undergoing bowel resection: results of a randomized trial of alvimopan and placebo with a standardized accelerated postoperative care pathway. Arch Surg. 2008 Nov;143(11):1098-105.
12. Buchler MW, Seiler CM, Monson JR, et al. Clinical trial: alvimopan for the management of post-operative ileus after abdominal surgery: results of an international randomized, double-blind, multicentre, placebo-controlled clinical study. Aliment Pharmacol Ther. 2008 Aug 1;28(3):312-25.
13. Harbaugh CM, Al-Holou SN, Bander TS, et al. A statewide, community-based assessment of alvimopan's effect on surgical outcomes. Ann Surg. 2013 Mar;257(3):427-32.
14. Delaney CP, Craver C, Gibbons MM, et al. Evaluation of clinical outcomes with alvimopan in clinical practice: a national matched-cohort study in patients undergoing bowel resection. Ann Surg. 2012 Apr;255(4):731-8.
15. Simorov A, Thompson J, Oleynikov D. Alvimopan reduces length of stay and costs in patients undergoing segmental colonic resections: results from multicenter national administrative database. Am J Surg. 2014 Sep 22;208(6):919-25.
16. Bell TJ, Poston SA, Kraft MD, et al.. Economic analysis of alvimopan in North American Phase III efficacy trials. Am J Health Syst Pharm. 2009 Aug 1;66(15):1362-8.
17. Flum DR, Fisher N, Thompson J, et al. Washington State's approach to variability in surgical processes/Outcomes: Surgical Clinical Outcomes Assessment Program (SCOAP). Surgery. 2005;138(5):821-8.
18. Devine E, Alfonso-Cristancho R, Devlin A, et al. A Model for Incorporating Patient and Stakeholder Voices in a Learning Healthcare Network: Washington State's Comparative Effectiveness Research Translation Network (CERTAIN). J Clin Epidemiol. 2013 April 2013;66:S122-9.
19. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. J Clin Epidemiol. 1992 Jun;45(6):613-9.
20. Washington state medicaid office of hospital finance rates and information [Internet]. [Accessed 7/21/15]. Available from: http://www.hca.wa.gov/medicaid/hospitalpymt.

21. Consumer Price Index [Internet]. [Accessed 7/21/25]. Available from: http://www.bls.gov/cpi/.
22. Moran JL, Solomon PJ, Peisach AR, Martin J. New models for old questions: generalized linear models for cost prediction. J Eval Clin Pract. 2007 June;13(3):381-9.
23. Singal AG, Higgins PD, Waljee AK. A primer on effectiveness and efficacy trials. Clin Transl Gastroenterol. 2014 Jan 2;5:e45.
24. Revicki DA, Frank L. Pharmacoeconomic evaluation in the real world. Effectiveness versus efficacy studies. Pharmacoeconomics. 1999 May;15(5):423-34.
25.. Ahmed Ali U, Dunne T, Gurland B, et al. Actual versus estimated length of stay after colorectal surgery: which factors influence a deviation? Am J Surg. 2014 Oct;208(4):663-9.
26.. Brooks JM, Ohsfeldt RL. Squeezing the balloon: propensity scores and unmeasured covariate balance. Health Serv Res. 2013 Aug;48(4):1487-507.
27. Poston S, Broder MS, Gibbons MM, et al. Impact of alvimopan (entereg) on hospital costs after bowel resection: results from a large inpatient database. P T. 2011 Apr;36(4):209-20.
28. Earnshaw SR, Kauf TL, McDade C, et al. Economic Impact of Alvimopan Considering Varying Definitions of Postoperative Ileus. J Am Coll Surg. 2015 Nov;221(5):941-50.
29. Touchette DR, Yang Y, Tiryaki F, Galanter WL. Economic analysis of alvimopan for prevention and management of postoperative ileus. Pharmacotherapy. 2012 Feb;32(2):120-8.
30. Lassen K, Soop M, Nygren J, et al. Consensus Review of Optimal Perioperative Care in Colorectal Surgery: Enhanced Recovery After Surgery (ERAS) Group Recommendations. Arch Surg. 2009 Oct;144(10):961-9.
31. Adam MA, Lee LM, Kim J, et al. Alvimopan Provides Additional Improvement in Outcomes and Cost Savings in Enhanced Recovery Colorectal Surgery. Ann Surg. 2015 Oct 22.
32. Weiss AJ, Elixhauser A. Trends in Operating Room Procedures in U.S. Hospitals, 2001-2011: Statistical Brief #171. Healthcare Cost and Utilization Project (HCUP) Statistical Briefs Rockville (MD); 2006.
33. Lee J. Rethinking spine care: Some health systems are moving beyond surgery in serving back pain patients. March 22, 2014; Available at: http://www.modernhealthcare.com/article/20140322/MAGAZINE/303229985. Accessed April 8, 2015.
34. Godil SS, Parker SL, O'Neill KR, et al. Comparative effectiveness and cost-benefit analysis of local application of vancomycin powder in posterior spinal fusion for spine trauma: clinical article. J Neurosurg Spine 2013; 19(7)550-555.
35. Calderone RR, Garland DE, Capen DA, Oster H. Cost of medical care for postoperative spinal infections. Orthop Clin North Am 1996 Jan;27(1):171-182.
36. Miller A. Hospital Reporting and "Never Events". Medicare Patient Management 2009(May/June):3/25/2015-http://www.medicarepatientmanagement.com/issues/04-03/mpmMJ09-NeverEvents.pdf.
37. Milstein A. Ending Extra Payments for "Never Events" - Stronger Incentives for Patients' Safety. The New England Journal of Medicine 2009;360:2388.
38. de Lissovoy G, Fraeman K, Hutchins V, et al. Surgical site infection: incidence and impact on hospital utilization and treatment costs. Am J Infect Control 2009 Jun;37(5):387-397.
39. Radcliff KE, Neusner AD, Millhouse PW, et al. What is new in the diagnosis and prevention of spine surgical site infections. Spine J 2015 Feb 1;15(2):336-347.
40. Edwards JR, Peterson KD, Mu Y, et al. National Healthcare Safety Network (NHSN) report: data summary for 2006 through 2008, issued December 2009. Am J Infect Control 2009, 37(10): 783-805.
41. Molinari RW, Khera OA, Molinari WJ,3rd. Prophylactic intraoperative powdered vancomycin and postoperative deep spinal wound infection: 1,512 consecutive surgical cases over a 6-year period. Eur Spine J 2012 Jun;21 Suppl 4:S476-82.
42. Rimoldi RL, Haye W. The use of antibiotics for wound prophylaxis in spinal surgery. Orthop Clin North Am 1996 Jan;27(1):47-52.
43. Tubaki VR, Rajasekaran S, Shetty AP. Effects of using intravenous antibiotic only versus local intrawound vancomycin antibiotic powder application in addition to intravenous antibiotics on

postoperative infection in spine surgery in 907 patients. Spine (Phila Pa 1976) 2013 Dec 1;38(25):2149-2155.
44. Bakhsheshian J, Dahdaleh NS, Lam SK, et al. The use of vancomycin powder in modern spine surgery: systematic review and meta-analysis of the clinical evidence. World Neurosurg 2014 Dec 19.
45. Evaniew N, Khan M, Drew B, et al. Intrawound vancomycin to prevent infections after spine surgery: a systematic review and meta-analysis. Eur Spine J 2015 Mar;24(3):533-542.
46. Khan NR, Thompson CJ, DeCuypere M, et al. A meta-analysis of spinal surgical site infection and vancomycin powder. J Neurosurg Spine 2014 Dec;21(6):974-983.
47. Viola RW, King HA, Adler SM, Wilson CB. Delayed infection after elective spinal instrumentation and fusion. A retrospective analysis of eight cases. Spine (Phila Pa 1976). 1997 Oct 15;22(20):2444,50; discussion 2450-1.
48. Bas T, Bas P, Blasco A, Bas JL. Chronic infections of the spine. Eur J Orthop Surg Traumatol. 2013 Jul;23 Suppl 1:S35-40.
49. Goodman SB, Yao Z, Keeney M, Yang F. The future of biologic coatings for orthopaedic implants. Biomaterials. 2013 Apr;34(13):3174-83.
50. Shaffer WO, Baisden JL, Fernand R, Matz PG, North American Spine Society. An evidence-based clinical guideline for antibiotic prophylaxis in spine surgery. Spine J 2013 Oct;13(10):1387-1392.
51. Lee MJ, Shonnard N, Farrokhi F, et al. The Spine Surgical Care and Outcomes Assessment Program (Spine SCOAP): a surgeon-led approach to quality and safety. Spine (Phila Pa 1976) 2015 Mar 1;40(5):332-341.
52. Cizik AM, Lee MJ, Martin BI, et al. Using the spine surgical invasiveness index to identify risk of surgical site infection: a multivariate analysis. J Bone Joint Surg Am 2012 Feb 15;94(4):335-342.
53. Mirza SK, Deyo RA, Heagerty PJ, et al. Towards standardized measurement of adverse events in spine surgery: conceptual model and pilot evaluation. BMC Musculoskelet Disord 2006 Jun 20;7:53.
54. Hu FB, Goldberg J, Hedeker D, et al. Comparison of population-averaged and subject-specific approaches for analyzing repeated binary outcomes. Am J Epidemiol 1998 Apr 1;147(7):694-703.
55. Austin PC, Mamdani MM, Stukel TA, et al. The use of the propensity score for estimating treatment effects: administrative versus clinical data. Stat Med 2005 May 30;24(10):1563-1578.
56. Weiss LE, Vaccaro AR, Scuderi G, et al. Pseudarthrosis After Postoperative Wound Infection in the Lumbar Spine. J Spinal Disord. 1997 Dec;10(6):482-7.
57. Schoenman J, Sutton J, Kintala S, Love D, Maw R. The Value of Hospital Discharge Databases. 2005. http://www.hcup-us.ahrq.gov/reports/final_report.pdf. Accessed 4/9/15.
59. Garrido MM, Kelley AS, Paris J, et al. Methods for Constructing and Assessing Propensity Scores. Health Serv Res. 2014;49(5): 1701-1720.
60. Brooks JM, Ohsfeldt RL. Squeezing the Balloon: Propensity Scores and Unmeasured Covariate Balance. Health Services Research 2013. 48;4:1487-1507.
61. Hubbard AE, Ahern J, Fleischer NL, et al. To GEE or Not to GEE: Comparing Population Average and Mixed Models for Estimating the Associations Between Neighborhood Risk Factors and Health. Epidemiology 2010;21:467-474.
62. Gardiner JC, Luo Z, Roman LA. Fixed effects, random effects, and GEE: What are the differences? Statist Med. 2009;28:221-239.