Evaluation of Legacy Patient Reported Outcome Measures as Performance Measures in Rehabilitation

Final Report

July 11, 2019

John M. Brooks

Cole G. Chapman

Adam Lutz

CER*ortho*
Center for Effectiveness Research
in Orthopaedics

Executive Summary

The Physical Therapy Industry Outcomes Workgroup contracted with the Center for Effectiveness Research in Orthopaedics (CERortho) to develop risk-adjusted performance measure algorithms for physical therapists using *legacy patient-reported outcome (PRO) measures and patient-level information* commonly available in electronic medical records (EMRs). The Workgroup asked for distinct performance measure algorithms for shoulder, knee, back, and neck, based on the quick Disability of Arm Shoulder Hand (qDASH),[1] Lower Extremity Functional Scale (LEFS),[2] Modified Low Back Disability Questionnaire (MDQ),[3] and Neck Disability Index (NDI)[4] PROs, respectively. Four physical therapy organizations contributed data for this study – WebPT, Intermountain, ATI, and PT Northwest.

The following steps were used to develop the risk-adjusted provider performance algorithms:

- Create a common analytical database of patient *episodes of care* necessary for risk-adjusted provider performance measurement that included legacy PRO measures and baseline patient factors that affect patient outcomes but are outside the control of the provider.
- Use the common analytical database to estimate separate regression models for shoulder, knee, back, and neck. Each model describes the average effect of each baseline patient factor on legacy outcome changes. For each patient these models provide a predicted average outcome change given the patient's baseline factors.
- The difference between each patient's actual outcome change and predicted outcome change is a risk-adjusted measure of provider performance for the patient episode.   A positive value means the provider performed better than average and a negative value means the provider performed below average.
- The average of the differences between actual and predicted patient outcomes across a provider's patient episodes is a risk-adjusted measure of provider performance.

Key findings:

- C*linically important improvements in PROs (median 8-16%) were observed across body regions over 12-14 visits*. High variability in average risk-adjusted PRO improvements exists across physical therapists.

- Baseline patient factors **currently available in existing physical therapy EMRs** for all data contributors explain a large portion of variation in PRO change within episodes. Payer type, patient socio-demographic factors, and comorbidities at baseline all had strong effects on PRO changes over episodes of care. Differences in these baseline patient factors must be accounted for to ensure fair performance comparisons of physical therapists

- Risk adjustment model specifications that included BMI and a comprehensive list of comorbidities increased explanatory power. Zip code level median income appears to capture socio-economic differences across patients.

- The completeness of data necessary for risk adjustment was a limitation of this project with only 8.8% of the patient episodes received had appropriated baseline and discharge PROs and only 6.3% had measures of the appropriate set of risk-adjustment factors. Physical therapy organizations must be committed institutional strategies that promote **the collection of PROs at baseline and discharge and baseline patient factors into existing EMRs**.

I. Background

The Physical Therapy Industry (PTI) Outcomes Registry Workgroup contracted with the Center for Effectiveness Research in Orthopaedics (CERortho) to evaluate the use of *legacy patient-reported outcome (PRO)* measures *and patient-level information* generally available in electronic medical records (EMRs) as the basis for comparative performance reporting of providers in physical therapy settings. CERortho was charged to develop distinct risk-adjusted provider performance models for shoulder, knee, back, and neck. Four distinct physical therapy providers contributed data for this study – WebPT, Intermountain, ATI, and PT Northwest.

For each provider, the average outcome change during their patients' episodes of care provides a measure of performance. However, some factors influencing patient outcome changes are outside the control of each provider and are unrelated to provider performance. To ensure that performance measures provide meaningful comparisons across providers, measures must *control* for baseline factors that affect outcome changes and are outside the control of the provider. "Risk-adjustment" regression models show, on average, how measured baseline patient factors affect outcome change. As a result, these risk-adjustment model estimates can predict the average outcome change over the care episode for each patient given each patient's measured baseline factors. The difference between a patient's actual outcome change and predicted outcome change is a risk-adjusted measure of provider performance for that patient. If this difference is positive, the patient received better than average care. If this difference is negative, the patient received less than average care. The average of these differences across a provider's patient episodes offers a risk-adjusted measure of provider performance. Risk-adjusted performance measures remain limited by the inability to measure the full set of factors that influence patient outcome changes and are outside the control of each provider.

CERortho assessed whether it is possible to create meaningful risk-adjustment models using legacy patient-reported outcome measures and baseline information available in EMRs for each patient. Table 1 in the request for proposal (RFP) contained the conceptual list of baseline factors the PTI Workgroup considered necessary for risk-adjustment. Risk-adjustment models can only be useful if these factors are measured and reported consistently across data contributors. The **objectives** of this research were to (1) create a common analytical database from the data supplied across contributors for performance measure development; (2) develop and evaluate the risk-adjusted performance measure algorithms for shoulder, knee, back, and neck found using this database.

II. Approach

CERortho performed the following steps to reach our objectives:

Step 1: Develop a set of *potential* risk-adjustment model specifications in consideration of the underlying concepts in Table 1 of the RFP. A patient episode can only be included in risk-adjustment model estimation if the episode provides changes in patient-reported outcome scores over the episode and all baseline patient factors specified in the model. If factors are unmeasured for some patients, tradeoffs will occur between the number of factors used in risk-adjustment modeling and the number of patients used to estimate these models. To assess the implications of these tradeoffs, CERotho created three distinct risk-adjustment model specifications. The "Optimal" specification contained the full range of factors in Table 1 of the RFP. Concepts in the Optimal specification included payer variables, height, weight, several health scores at index, sex, age, race, median income based on patient residence Zip

3

Code and comorbidities. "Minimal" contained the subset of factors most consistently measured across patients in the data, and "Practical" which added BMI and additional comorbidities to the minimal specification.

Step 2: Assess the ability of the data supplied from each contributor to measure the factors listed in each risk-adjustment model specification. Each data contributor agreed to send CERortho a data set organized at the "episode of care" level. Each observation at the episode of care level should contain baseline patient factors and changes in outcome scores over the episode. We mapped the measures provided by each data contributor into the factors in the risk-adjustment model specifications listed in Step 1. A patient was defined as "***available"*** for model estimation if all factors required for that specification were measured for that patient. We assessed the number of patients available from each data contributor for each risk-adjustment model specification. We also created summary statistics for each factor in the risk-adjustment models for each body region (shoulder, neck, knee, back) and model specification.

Step 3: Estimate risk-adjustment model specifications by body region (shoulder, neck, knee, back). Estimate risk-adjustment models for each body region for each model specification listed in Step 2. Assess the overall ability of risk-adjustment modeling to control for baseline patient characteristics that are outside the control of the provider.

Step 4: Select risk-adjustment model factors based on the tradeoffs between the number of factors specified and available sample size. For a patient's episode to be included in a regression model outcome change and all specified risk-adjustment factors must be measured. Because risk-adjustment factors were unmeasured in many of the episodes of care provided the data contributors, the number of patients *available* for estimation diminishes as additional factors are specified. In this step, we contrasted the number of episodes available for estimation across the Minimal, Practical, and Optimal specifications with the information provided each factor to choose risk-adjustment specifications for each body region.

Step 5: Assess alternative functional forms for the risk-adjustment specification selected. The risk-adjustment model specifications selected in Step 4 were in terms of the baseline factors specified. Given the factors specified, in this step we assessed whether specifying these factors as non-linear functions of these factors improved the ability of the models to predict outcome change.

Step 6: Demonstrate the variation in physical therapist performance generated by this approach. For risk-adjustment model specification and functional forms selected in Step 5, we assessed the distribution of risk-adjusted provider performance measures generated by these models in comparison to the minimal clinically important difference (MCID) in the legacy outcome scores by body region (6 - back,[3] 10 - neck,[5] 11 - shoulder,[6] 9 - knee,[2]).

III. Results

Baseline Risk-Adjustment Factor Measures

Table 1 lists the baseline factors based listed in the RFP for risk-adjustment and which of these factors were included in each of model specification. The left most column of Table 1 contains the factors from the RFP and the second column describes how each factor was measured for each episode of care. The last three columns indicate the factors specified in the "Minimal", "Practical", and "Optimal" risk-

adjustment model specifications. Most of the measured factors in Table 1 are *binary variables* in which "1" indicates the factor existed for the patient at baseline, "0" if it did not. For example, the factor History_Smoking equals "1" if a patient stated a history of smoking at her/his initial physical therapy visit. Patients without a history of smoking would have a value of "0" for this factor. Some factors were specified as *continuous* at baseline including Height_CM, Weight_Lbs, BMI, Outcome_Index, Pain_Score_Index, General_Health_Mental_Index, General_Health_Physical_Index, and Patient Zip Code Median Income. Patient Zip Code Median Income data were obtained from US Census Bureau by CERortho and assigned to each patient based on reported residence Zip Code.[7]

Patient Episodes *Available* for Analyses

Figure 1 provides the number of the unique patient episodes received from the four data contributors and the number of episodes available for use in risk-adjustment modeling by body region for each model specification. We defined an episode from a data contributor as ***available*** for a model specification if all required factors for that specification were measured in that episode. The first box of Figure 1 shows the total number of episodes of care received across contributors. CERortho received data on over 4 million unique episodes. Of that number, 520,847 episodes had both index and discharge patient-reported outcomes for the legacy outcome measures listed in the request for proposal (RFP). The RFP listed the Modified Low Back Disability Questionnaire (MDQ) for back, Neck Disability Index (NDI), Disability of Arm Shoulder Hand (Dash) for shoulder, and the Lower Extremity Functional Scale (LEFS) for knee. CERortho will supply information back to each data contributor individually describing data quality. One data contributor did not supply patient-reported outcomes for any episode of care. A second data contributor had only 16% of episodes with an outcome score at the initial visit and 0.21% of episodes with an outcome score at discharge. A third contributor had the highest percentage of outcome scores available at initial visit (64%) and discharge (64%). This contributor provided a higher percentage of episodes with complete outcome information but used outcomes from the Knee Outcome Survey (KOS) for knee instead of LEFS. Staying consistent with the protocol, our *primary analysis* only included episodes that used the LEFS. However, we performed a secondary analysis for knee that included episodes with either KOS or LEFS outcomes standardized on a 0-100 scale and the regression estimates were consistent. Next, 382,601 episodes had outcome scores consistent with the body region listed by the data contributor for the episode. Lastly, episodes were available for regression analysis if all risk-adjustment factors included in the respective specification were measured. The Minimal Model Specification had 315,210 episodes (7.2% of total) available for analysis (82,801 shoulder, 66,977 neck, 13,203 knee, 152,229 back). The Practical Model Specification had 272,933 episodes (6.3% of total) available for analysis (71,662 shoulder, 58,164 neck, 10,892 knee, 132,215 back). Whereas, the Optimal Model Specification had only 4,926 episodes (0.1% of total) available for analysis (1,148 shoulder, 1,123 neck, 188 knee, 2,467 back). The large drop-off of available episodes from the Practical to the Optimal model specification was due mainly to missing Race variables and alternative baseline measures of health.

Characteristics of Patients Available for Risk-Adjustment Modeling

Tables 3.1, 3.2, 3.3, and 3.4 contain summary statistics for the patient episodes available for risk-adjustment modeling by model specification for shoulder, neck, knee, and back, respectively. Each table contains estimates for the three specifications: Minimal, Practical, and Optimal.   The patient-reported outcome scores for each body region were normalize to a 0 (worst) to 100 (best) scale. The average

outcome scores at index (found within the Health Status at Index section of these tables) was around 55 for shoulder, 65 for neck, 44 for knee, and 64 for back. The large interquartile ranges for these index outcome scores indicate a broad span of initial scores for the patients in our sample reflecting large differences in the initial conditions of patients in the sample. Most patients in the sample had commercial insurance with Medicare second. Interestingly, "Auto" was the third highest for neck injuries. Most of the patients in the sample had chronic conditions and most initiated treatment without prior surgery. More than half of the patients were female with neck injuries having the largest female percentage. The highest reported comorbidities were Arthritis, Breathing Difficulties, Diabetes, Psychological Conditions, and Smoking. Patient median age was in the low to mid-50s with 50% of the patients between the late 30s to late 60s. Zip Code-level median incomes were around $65,000 with interquartile ranges from 50 to 75 thousand. Median BMI was approximately 28 with interquartile ranges approximately 24-30.

Change in the patient-reported outcome score across an episode of care was the dependent variable in our risk-adjustment regression models. The median outcome score changes for both shoulder and knee were around 16 with wide interquartile ranges around these values. The median change in the neck score was around 10 and the median change in the back score was approximately 8. Likewise, there was substantial variation in these change scores across episodes as reflected by their respective interquartile ranges. Patients in our sample had a substantial number of physical therapy visits within their episodes of care with median shoulder, neck, knee, and back visits of around 13, 12, 14, and 12, respectively.

Risk-Adjustment Model Estimates

Tables 4.1, 4.2, 4.3, and 4.4 contain the risk-adjustment model regression estimates for shoulder, neck, knee, and back, respectively. Each table contains the estimates from three specifications: Minimal, Practical, and Optimal.   The observations in these analyses are single episodes of care as listed in Figure 1. The outcome variables for each body region were scaled from 0-100 with 0 being the lowest outcome and 100 the best. The dependent variable in each specification is the change in outcome score for each patient from the episode index visit through episode discharge (e.g. 75-55 = 20). In addition, because the regression models are linear, the coefficient estimates for all risk-adjustment variables have direct interpretations. For factors specified as binary variables, the estimated coefficient equals the expected difference in the outcome change score between patients with that condition at baseline and those who did not. For example, in the shoulder practical specification in Table 4.1, patients with commercial insurance had on average a 4.64 higher outcome change over their care episode than patients with auto insurance (the reference group). In contrast, patients with workman's compensation insurance had on average a 5.51 lower outcome change over their care episode than patients with auto insurance. For continuous factors, the estimated coefficients reflect the difference in expected outcome change for a one unit change in the factor. For example, in the practical shoulder specification, a one unit change in BMI (e.g. 27 to 28) was associated with a 0.05 unit drop in outcome change over the episode of care. In contrast, we measured Zip Code median income in $10,000 increments. Therefore, $10,000 higher Zip Code median income is associated with a .24 unit increase in the average outcome change over the episode of care.

The R^2 statistics for each specification provide an estimate of the percentage of variation in PRO change scores is attributable to the set of risk-adjustment factors specified in each model. In the practical specifications, the specified risk-adjustment factors explain 40% of the PRO change variation in shoulder,

17% in neck, 40% in knee, and 17% in back. The F-statistics found on the last row of each column in Tables 4.1, 4.2, 4.3, and 4.4 are tests indicating whether the set of baseline factors specified in each model describe a significant portion of the variation in outcome change. In the twelve regression models, these F-statistics were large and statistically significant indicating the importance of these baseline factors on outcome changes over the episode of care. Since these factors our outside the control of the provider, these results suggest that risk-adjustment is necessary to yield performance measures that can provide meaningful comparisons across providers. For performance measurement purposes, it is assumed that the portion of the variation in PRO change remaining after controlling for the measured set of baseline factors is attributable to care process choices made by the provider. The validity of these measures relies on this assumption.

Table 5 contains the partial F-statistics that describe the the relative contributions of sets of baseline factors in describing outcome change over episodes of care. For three of the four body regions, the set of Payer variables have strongest relationships with outcome changes over the episode. However, each set of baseline factors had a strong independent influence on outcome change.

Preliminary Model Specification Choice

Relative to the Practical specification, the Optimal specification includes Race variables, and three health outcome measures at baseline beyond the index outcome score (Pain, Mental Health, Physical Health). These variables were the highest percentage missing across the risk-adjustment factors and are the main cause that so few patient episodes (0.1%) were "available" for the Optimal specification regressions. Given the information presently available across data contributors, it does not seem feasible to include Race and these additional baseline health scores in the risk-adjustment models. Fortunately, these factors did not explain a significant portion of the change in outcome scores in the Optimal regression models. It appears that the Outcome Score at Index and the Patient Zip Code Median Income capture the baseline differences in the condition of the relevant body region and socio-demographics, respectively.

Relative to the Minimal specification, the Practical Specification included BMI and several additional baseline comorbidities. Because BMI was missing for around 14% of the patients, the sample sizes for the Practical Specifications were lower than the Minimal specification for each body region.   However, baseline BMI had a statistically significant effect on outcome change scores for three of the four body regions. In addition, many of the added comorbidities also affected the outcome change scores. As a result, we believe the Practical Specification provides a strong foundation of risk-adjustment modeling given the information presently collected by physical therapy practices.

Functional Form Assessment

The Practical Model specifications found in Tables 4.1-4.4 are all linear in nature. Linear specifications restrict the effect of each baseline factor on PRO change to be constant across patients and these effects do not vary with the initial level of the factor or the levels of other factors in the model. If the relationships between PRO changes and measured baseline factors vary with the initial value of the baseline factor or with the values of the other baseline factors, linear models would not sufficiently risk-adjust. For example, in the Practical shoulder model in Table 4.1, the estimated coefficient for BMI is -0.05. This linear specification suggests that a one unit increase in baseline BMI is associated with a 0.05 decrease in outcome change. This decrease is the same no matter if a patient is at a low BMI at baseline

(e.g. 21) or a high BMI (e.g. 30) at baseline. In addition, the relationship between baseline BMI and outcome change is the same for males and females, young and old, etc. To evaluate these restrictions, each Practical Model specification was also estimated with a series of nonlinear terms to assess whether these terms contributed to the ability of the baseline factors to describe variation in PRO change across patients. These terms included "interactions" across factors (e.g. Age*Chronic) and terms that "squared" each of the continuous factors (e.g. BMI*BMI). We assessed whether adding these terms into the regression models lead to a significantly higher portion of the variation in PRO change explained relative to the linear model. Across all four body regions, the nonlinear terms as a group added little to each model's ability to describe outcome change. As a result, CERortho favors using the linear Practical Model specifications found in Tables 4.1-4.4 for the basis of physical therapist risk-adjustment.

Performance Measure Distributions

Figures 2, 3, 4, and 5 show the distribution of performance measures across physical therapists for shoulder, neck, knee, and back patients, respectively. Each figure contains performance measures for only the physical therapists in the database who treated more than 40 patients in the respective body region. The Y-axis represents the average of difference in actual versus predicted outcome change scores and each observation is the average value for a single physical therapist. Physical therapists are distributed along the X-axis from lowest to highest performance measures. Physical therapists with a performance measure greater than zero on the Y axis performed better than average given the characteristics of their patients. The opposite is true of the physical therapists with a performance measure below zero on the Y axis.

The figures also contain horizontal lines representing the minimal clinically important difference (MCID) for each respective outcome score above and below average (6 -back,[3] 10 - neck,[5] 11 - shoulder,[6] 9 - knee,[2]). Physical therapists below the lower line have patient outcome changes more than the MCID below the average score. Whereas, physical therapists above the upper line have patient outcome changes more than the MCID above the average score. Most of physical therapists in our database had performance measures within the MCID range of average. For knee and shoulder all the physical therapists (knee 38; shoulder 607) included in the analysis had performance measures within plus or minus the MCID of the average patient. For back (1,213 physical therapists), nine had performance measures below the minus MCID threshold and 14 had performance measures above the plus MCID threshold. For neck (390 physical therapists), one had performance measures below the minus MCID threshold and one had performance measures above the plus MCID threshold.

IV. Limitations

The performance measurement approach attributes all change in patient outcomes not attributable to baseline patient factors to provider performance. Other unmeasured baseline patient factors may still exist that affect patient outcomes. As a result, the measures calculated here will be a biased measure of a physical therapist's performance if the therapist cares for patients with a disproportionate distribution of these unmeasured factors. In addition, a significant limitation of this study was the low number of patients who had complete data. Patient outcomes with the patients who had complete data could look different than those with incomplete data and these differences are not reflected in the risk-adjustment models estimated here.. There also exists variation with both the outcome and risk adjustment data collected by the contributors to this study. Standardization of the data collected in this industry is needed.

V. Summary Findings

This study shows that risk-adjusted adjusted performance measurement is feasible using *legacy patient-reported outcome (PRO) measures and patient-level information* that is collectable from existing electronic medical record (EMR) systems. Baseline factors outside the control of providers explain a substantial portion of the variation in episode-level outcome changes across patients. As result, risk-adjustment for these factors is necessary for meaningful performance comparisons across providers. *Given current data collection practices*, the "Practical" model specification offers a strong alternative for use in performance measurement.   Payer status, patient socioeconomic characteristics, and comorbidities at care initiation all had significant effects on subsequent outcome changes and should be "risk-adjusted" to ensure that performance measures provide fair comparisons across physical therapists. Physical therapy firms must be committed to the collection of baseline patient information and patient-reported outcomes across each episode of care to utilize this approach.

Table 1: Model Factors Included Across Risk-Adjustment Model Specifications				
Factors	**Definition**	**Minimal**	**Practical**	**Optimal**
Patient-Reported Outcome Change	(patient-reported outcome score at episode discharge) – (patient-reported outcome score at episode index visit). All outcome scores scaled from 0 (worst) to 100 (best).	X	X	X
Payer Indicator Variables		X	X	X
Auto (reference group)	1 if patient insured through Auto insurance, 0 otherwise.	X	X	X
Commercial	1 if patient insured through Commercial insurance, 0 otherwise.	X	X	X
Medicaid	1 if patient insured through Medicaid insurance, 0 otherwise.	X	X	X
Medicare	1 if patient insured through Medicaid insurance, 0 otherwise.	X	X	X
Other	1 if patient insured through NED DEF insurance, 0 otherwise.	X	X	X
Workman's Comp	1 if patient insured through Workman's Compensation, 0 otherwise.	X	X	X
Outcome Score at Index	Patient index outcome score (0-100)	X	X	X
Health Status at Index				
Height_CM	Patient height in centimeters			X
Weight_Lbs	Patient weight in pounds			X
BMI	Patient BMI score		X	
Acute (ref))	1 if patient injury is acute, 0 otherwise.	X	X	X
Chronic	1 if patient injury is chronic, 0 otherwise.	X	X	X
First PT Visit Before Surgery (ref)	1 if patient index visit occurred before surgery, 0 otherwise.	X	X	X
First PT Visit After Surgery	1 if patient index visit occurred after surgery, 0 otherwise.	X	X	X

Pain_Score_Index	Patient index pain score			X
General_Health_Mental_Index	Patient index mental health score			X
General_Health_Physical_Index	Patient index physical health score			X
Socio-Demographics				
Female (ref)	1 if patient is female, 0 otherwise.	X	X	X
Male	1 if patient is male, 0 otherwise.	X	X	X
Age	Age of patient at index visit	X	X	X
Race				X
Asian/Pacific (ref)	1 if patient is Asian/Pacific origin, 0 otherwise.			X
African-American	1 if patient is African-American, 0 otherwise.			X
Caucasian	1 if patient is Caucasian, 0 otherwise.			X
Eskimo/American Indian	1 if patient is Eskimo/American Indian, 0 otherwise.			X
Hispanic	1 if patient is Hispanic, 0 otherwise.			X
Patient Zip Code Median Income	Median income in patient residence Zip Code in 10K (e.g. 55,000 specified as 5.5)	X	X	X
Comorbidities at Index				
History_Arthritis	1 if patient has a history of arthritis, 0 otherwise.		X	X
History_Blood_Clot	1 if patient has a history of blood clots, 0 otherwise.		X	X
History_Breathing_Difficulties_Asthma	1 if patient has a history of breathing difficulties, 0 otherwise.		X	X
History_Cancer	1 if patient has a history of cancer, 0 otherwise.	X	X	X
History_Diabetes	1 if patient has a history of diabetes, 0 otherwise.	X	X	X
History_Kidney_Condition	1 if patient has a history of kidney conditions, 0 otherwise.		X	X
History_Osteoporosis	1 if patient has a history of osteoporosis, 0 otherwise.		X	X
History_Pacemaker	1 if patient has a pacemaker, 0 otherwise.		X	X
History_Psychological_Condition	1 if patient has a history of psychological conditions, 0 otherwise.		X	X
History_Seizures	1 if patient has a history of seizures, 0 otherwise.		X	X
History_Smoking	1 if patient has a history of smoking, 0 otherwise.	X	X	X
History_Stroke	1 if patient had a previous stroke, 0 otherwise.		X	X
History_Unexplained_Weight_Loss	1 if patient had a previous unexplained weight loss, 0 otherwise.	X	X	X

Figure 1 Sample Size Summary for Risk-Adjustment Analysis By Body Region and Model Specification

Optimal Specification
- shoulder 1,148
- neck 1,123
- knee 188
- back 2,467

Practical Specification
- shoulder 71,662
- neck 58,164
- knee 10,892
- back 132,215

Minimal Specification
- shoulder 82,801
- neck 66,977
- knee 13,203
- back 152,229

Total Episodes Provided
4,312,909

Episodes With Valid Outcome Scores For Selected Body Regions
520,847

Episodes With Valid Outcome Score that Matches Designated Body Part
382,601

Episodes With Complete Information For Risk-Adjustment Analysis

Table 3.1: Descriptive Information for **Shoulder Patients** by Model Specification			
	Median (Interquartile Range) or % (Count) by Model Specification		
	Minimal	*Practical*	*Optimal*
Episode Count	82,801	71,662	1,148
PRO Change	16 (2, 31)	16 (2, 31)	18 (3, 34)
Visit Count	13 (7, 20)	13 (7, 20)	17 (8, 26)
Payer Variables (ref = auto)			
Commercial	58% (47901)	57% (40966)	40% (457)
Medicaid	5% (4126)	5% (3566)	6% (71)
Medicare	21% (17779)	22% (15961)	13% (146)
Other	2% (1705)	2% (1353)	2% (22)
Workman's Comp	10% (8732)	10% (7604)	33% (373)
Auto	3% (2558)	3% (2212)	7% (79)
Outcome Score at Index	55(36,73)	55 (36,73)	43 (25,61)
Health Status at Index			
Height_CM	*Not in Model*	*Not in Model*	167.6 (160.0,175.2)
Weight_Lbs	*Not in Model*	*Not in Model*	178 (150.5,205.5)
BMI	*Not in Model*	27.5 (23.7-31.3)	*Not in Model*
Chronic	84% (69814)	85% (60667)	*Not in Model*
First PT Visit After Surgery (ref = before)	27% (22469)	27% (19477)	34% (388)
Pain Score at Index	*Not in Model*	*Not in Model*	2 (2,2)
General Mental Score at Index	*Not in Model*	*Not in Model*	39.7 (34.5,44.9)
General Physical Score at Index	*Not in Model*	*Not in Model*	37.3 (32.9,41.7)

Socio-demographics			
Male (ref = female)	48% (40178)	48% (34271)	45% (512)
Age	55.3 (43,65.9)	55.8 (43.5,66.5)	53 (42.5,63.5)
Patient Zip Code Median Income ($10,000)	6.44 (4.99-7.89)	6.44 (4.99-7.89)	5.75 (4.27,7.23)
African-American	*Not in Model*	*Not in Model*	9% (101)
Caucasian	*Not in Model*	*Not in Model*	44% (505)
Eskimo/American Indian	*Not in Model*	*Not in Model*	3% (34)
Hispanic	*Not in Model*	*Not in Model*	1% (12)
Asian/Pacific (ref)	*Not in Model*	*Not in Model*	43% (496)
Comorbidities at Index			
History_Arthritis	*Not in Model*	38% (27076)	31% (360)
History_Blood_Clot	*Not in Model*	4% (2876)	3% (29)
History_Breathing_Difficulties_Asthma	*Not in Model*	17% (12140)	13% (154)
History_Cancer	8% (6983)	9% (6293)	5% (59)
History_Diabetes	14% (11911)	15% (10692)	20% (224)
History_Kidney_Condition	*Not in Model*	7% (4848)	3% (29)
History_Osteoporosis	*Not in Model*	9% (6304)	6% (66)
History_Pacemaker	*Not in Model*	1% (759)	0% (4)
History_Psychological_Condition	*Not in Model*	13% (9516)	3% (38)
History_Seizures	*Not in Model*	2% (1183)	1% (16)
History_Smoking	13% (10503)	13% (9409)	12% (136)
History_Stroke	*Not in Model*	2% (1632)	2% (26)
History_Unexplained_Weight_Loss	1% (1195)	2% (1095)	1% (16)

	Median (Interquartile Range) or % (Count) by Model Specification		
	Minimal	*Practical*	*Optimal*
Episode Count	**66,977**	**58,164**	**1,123**
PRO Change	**10 (0, 20)**	**10 (0, 20)**	**10 (-1,21)**
Visit Count	**12 (8, 17)**	**12 (8, 17)**	**14 (9,19)**
Payer Variables (ref = auto)			
Commercial	48% (32465)	48% (27859)	34% (386)
Medicaid	8% (5251)	8% (4595)	9% (98)
Medicare	18% (11936)	19% (10784)	13% (149)
Other	2% (1053)	1% (872)	1% (12)
Workman's Comp	6% (4100)	6% (3586)	14% (152)
Auto	18% (12172)	18% (10468)	29% (326)
Outcome Score at Index	66 (53,79)	64 (51,77)	60 (45,75)
Health Status at Index			
Height_CM	*Not in Model*	*Not in Model*	165.1 (158.7,171.5)
Weight_Lbs	*Not in Model*	*Not in Model*	175 (147.5,202.5)
BMI	*Not in Model*	27.1 (23.4,30.8)	*Not in Model*
Chronic	90% (60252)	90% (52399)	*Not in Model*
First PT Visit After Surgery (ref = before)	5% (3676)	6% (3251)	7% (83)
Pain Score at Index	*Not in Model*	*Not in Model*	2 (2,2)
General Mental Score at Index	*Not in Model*	*Not in Model*	38.9 (33.7,44.1)

Table 3.2: Descriptive Information for **Neck Patients** by Model Specification

General Physical Score at Index	*Not in Model*	*Not in Model*	36.9 (32.8,41.0)
Socio-demographics			
Male (ref = female)	36% (23966)	36% (20829)	34% (379)
Age	52.3 (40.3,64.3)	52.7 (40.7,64.7)	50.2 (39.1,61.3)
Patient Zip Code Median Income ($10,000)	6.35 (4.92,7.78)	6.32 (4.89,7.75)	5.53 (4.02,7.05)
African-American	*Not in Model*	*Not in Model*	10% (107)
Caucasian	*Not in Model*	*Not in Model*	52% (589)
Eskimo/American Indian	*Not in Model*	*Not in Model*	2% (25)
Hispanic	*Not in Model*	*Not in Model*	1% (11)
Asian/Pacific (ref)	*Not in Model*	*Not in Model*	10% (107)
Comorbidities at Index			
History_Arthritis	*Not in Model*	38% (22114)	33% (366)
History_Blood_Clot	*Not in Model*	4% (2195)	2% (26)
History_Breathing_Difficulties_Asthma	*Not in Model*	19% (11003)	15% (168)
History_Cancer	8% (5130)	8% (4581)	5% (55)
History_Diabetes	12% (8025)	12% (7142)	14% (162)
History_Kidney_Condition	*Not in Model*	6% (3708)	3% (32)
History_Osteoporosis	*Not in Model*	9% (5272)	5% (60)
History_Pacemaker	*Not in Model*	1% (632)	0% (5)
History_Psychological_Condition	*Not in Model*	15% (8902)	4% (44)
History_Seizures	*Not in Model*	2% (1198)	1% (11)
History_Smoking	14% (9650)	15% (8606)	11% (127)
History_Stroke	*Not in Model*	2% (1405)	3% (29)

History_Unexplained_Weight_Loss	2% (1497)	2% (1365)	1% (16)

Table 3.3: Descriptive Information for **Knee Patients** by Model Specification			
	Median (Interquartile Range) or % (Count) by Model Specification		
	Minimal	*Practical*	*Optimal*
Episode Count	13,203	10,892	188
PRO Change	16 (2, 31)	16 (2, 31)	14 (-2, 30)
Visit Count	14 (8, 20)	14 (9, 20)	16 (10, 23)
Payer Variables (ref = auto)			
Commercial	57% (7536)	56% (6081)	51% (96)
Medicaid	7% (981)	7% (792)	8% (15)
Medicare	24% (3226)	26% (2815)	18% (34)
Other	2% (262)	2% (203)	2% (3)
Workman's Comp	7% (912)	7% (756)	16% (31)
Auto	2% (286)	2% (245)	5% (9)
Outcome Score at Index	44 (26,62)	44 (25,63)	43.8 (23.8,63.8)
Health Status at Index			
Height_CM	*Not in Model*	*Not in Model*	165.1 (157.5,172.7)
Weight_Lbs	*Not in Model*	*Not in Model*	175.5 (148.4,202.6)
BMI	*Not in Model*	27.8 (23.6,32)	*Not in Model*
Chronic	99.6% (13149)	100% (10892)	*Not in Model*
First PT Visit After Surgery (ref = before)	33% (4378)	33% (3604)	33% (62)
Pain Score at Index	*Not in Model*	*Not in Model*	2 (2,2)
General Mental Score at Index	*Not in Model*	*Not in Model*	40.2 (35.5,44.9)
General Physical Score at Index	*Not in Model*	*Not in Model*	37.7 (33.7,41.7)

Socio-demographics			
Male (ref = female)	40% (13540)	43% (4631)	45% (84)
Age	56.1 (30,72.2)	56.9 (44.4,72.4)	52.4 (38.2,66.6)
Patient Zip Code Median Income ($10,000)	6.25 (4.88,7.62)	6.21 (4.77,7.65)	5.17 (3.72,6.62)
African-American	*Not in Model*	*Not in Model*	10% (19)
Caucasian	*Not in Model*	*Not in Model*	54% (102)
Eskimo/American Indian	*Not in Model*	*Not in Model*	2% (3)
Hispanic	*Not in Model*	*Not in Model*	2% (3)
Asian/Pacific (ref)	*Not in Model*	*Not in Model*	10% (19)
Comorbidities at Index			
History_Arthritis	*Not in Model*	41% (4475)	36% (67)
History_Blood_Clot	*Not in Model*	5% (522)	2% (4)
History_Breathing_Difficulties_Asthma	*Not in Model*	14% (1576)	12% (22)
History_Cancer	9% (1184)	9% (1012)	4% (7)
History_Diabetes	13% (1708)	13% (1448)	13% (24)
History_Kidney_Condition	*Not in Model*	4% (440)	2% (3)
History_Osteoporosis	*Not in Model*	10% (1037)	10% (18)
History_Pacemaker	*Not in Model*	1% (79)	1% (1)
History_Psychological_Condition	*Not in Model*	5% (526)	1% (2)
History_Seizures	*Not in Model*	2% (165)	1% (1)
History_Smoking	10% (1342)	10% (1128)	13% (25)
History_Stroke	*Not in Model*	3% (374)	4% (7)
History_Unexplained_Weight_Loss	1% (164)	1% (134)	2% (3)

Table 3.4: Descriptive Information for **Back Patients** by Model Specification			
	Median (Interquartile Range) or % (Count) by Model Specification		
	Minimal	*Practical*	*Optimal*
Episode Count	152,229	132,215	2,467
PRO Change	8 (-1, 17)	8 (-1, 17)	8 (-1, 17)
Visit Count	12 (8, 17)	12 (8, 17)	13 (9, 18)
Payer Variables (ref = auto)			
Commercial	51% (76995)	50% (65738)	38% (937)
Medicaid	9% (14208)	9% (12367)	12% (288)
Medicare	25% (38664)	26% (34795)	17% (415)
Other	2% (2672)	2% (2164)	2% (48)
Workman's Comp	7% (10700)	7% (9328)	19% (467)
Auto	6% (8990)	6% (7823)	13% (312)
Outcome Score at Index	64 (51,77)	64 (51,77)	58 (45,71)
Health Status at Index			
Height_CM	*Not in Model*	*Not in Model*	167.6 (160.0,175.2)
Weight_Lbs	*Not in Model*	*Not in Model*	182 (152.5,211.5)
BMI	*Not in Model*	28.2 (24.0,32.4)	*Not in Model*
Chronic	89% (135036)	89% (117789)	*Not in Model*
First PT Visit After Surgery (ref = before)	7% (10421)	7% (9174)	8% (203)
Pain Score at Index	*Not in Model*	*Not in Model*	2 (2,2)
General Mental Score at Index	*Not in Model*	*Not in Model*	38.1 (33.1,43.1)
General Physical Score at Index	*Not in Model*	*Not in Model*	37.0 (33.0,41.0)

Socio-demographics			
Male (ref = female)	42% (64278)	42% (55820)	43% (1062)
Age	55 (40.5,69.6)	55.6 (41.5,69.7)	52.3 (40.2,64.4)
Patient Zip Code Median Income ($10,000)	6.26 (4.83,7.69)	6.23 (4.82,7.64)	5.29 (3.82,6.76)
African-American	*Not in Model*	*Not in Model*	8% (205)
Caucasian	*Not in Model*	*Not in Model*	52% (1281)
Eskimo/American Indian	*Not in Model*	*Not in Model*	2% (54)
Hispanic	*Not in Model*	*Not in Model*	2% (41)
Asian/Pacific (ref)	*Not in Model*	*Not in Model*	36% (886)
Comorbidities at Index			
History_Arthritis	*Not in Model*	42% (55260)	36% (894)
History_Blood_Clot	*Not in Model*	4% (5804)	3% (72)
History_Breathing_Difficulties_Asthma	*Not in Model*	20% (25787)	15% (370)
History_Cancer	9% (13364)	9% (12003)	4% (103)
History_Diabetes	15% (22120)	15% (19822)	19% (476)
History_Kidney_Condition	*Not in Model*	8% (10339)	4% (98)
History_Osteoporosis	*Not in Model*	11% (14700)	8% (197)
History_Pacemaker	*Not in Model*	2% (1958)	0% (12)
History_Psychological_Condition	*Not in Model*	15% (19627)	5% (128)
History_Seizures	*Not in Model*	2% (2589)	2% (40)
History_Smoking	15% (23351)	16% (20829)	14% (348)
History_Stroke	*Not in Model*	3% (3783)	3% (81)
History_Unexplained_Weight_Loss	2% (3252)	2% (2972)	3% (77)

Table 4.1: Risk-Adjustment Regression Models for **Shoulder Patients** by Specification			
	Dependent Variable		
	Patient-Reported Outcome Change over Care Episode		
Model Specification	minimal	practical	optimal
Patient-Reported Outcome Change (Median, Interquartile Range)	16 (2, 31)	16 (2, 31)	18 (3, 34)
Visit Count (Median, Interquartile Range)	13 (7, 20)	13 (7, 20)	17 (8, 26)
	Estimate (95% confidence interval)	Estimate (95% confidence interval)	Estimate (95% confidence interval)
Payer Variables (ref = auto)			
Commercial	4.71*** (4.05, 5.37)	4.64*** (3.93, 5.35)	0.84 (-3.69, 5.36)
Medicaid	-1.98*** (-2.80, -1.16)	-1.88*** (-2.76, -1.01)	-3.80 (-9.77, 2.17)
Medicare	4.65*** (3.94, 5.36)	4.94*** (4.17, 5.71)	2.01 (-3.42, 7.44)
Other	2.52*** (1.60, 3.64)	2.78*** (1.66, 3.90)	-5.81 (-14.65, 3.03)
Workman's Comp	-4.92*** (-5.65, -4.19)	-5.51*** (-6.30, -4.73)	-9.99*** (-14.61, -5.37)
Outcome Score at Index	-0.55*** (-0.56, -0.54)	-0.56*** (-0.57, -0.55)	-0.50*** (-0.56, -0.45)
Health Status at Index			
Height_CM			-0.05 (-0.19, 0.10)
Weight_Lbs			0.02 (-0.01, 0.05)
BMI		-0.05*** (-0.07, -0.03)	
Chronic (ref = acute)	-0.47** (-0.79, 0.15)	-1.84*** (-2.21, -1.46)	no variation in data
First PT Visit After Surgery (ref = before)	6.33*** (6.04, 6.62)	6.14*** (5.83, 6.44)	3.76*** (1.19, 6.33)
Pain Score at Index			-0.31 (-1.71, 1.10)
General Mental Score at Index			0.11 (-0.04, 0.25)
General Physical Score at Index			0.07 (-0.09, 0.24)
Socio-demographics			

Male (ref = female)	2.06*** (1.83, 2.30)	1.70*** (1.45, 1.96)	1.23 (-1.81, 4.28)
Age	-0.11*** (-0.12, -0.10)	-0.09*** (-0.10, -0.08)	-0.16*** (-0.25, -0.07)
Patient Zip Code Median Income	0.25*** (0.20, 0.30)	0.24*** (0.18, 0.29)	0.44* (-0.06, 0.94)
African-American (ref = Asian/Pacific)			0.90 (-3.27, 5.06)
Caucasian (ref = Asian/Pacific)			-0.06 (-2.65, 2.54)
Eskimo/American Indian (ref = Asian/Pacific)			1.80 (-4.45, 8.51)
Hispanic (ref = Asian/Pacific)			-5.04 (-15.59, 5.50)
Comorbidities at Index			
History_Arthritis		-1.09*** (-1.38, -0.80)	-0.93 (-3.74, 1.87)
History_Blood_Clot		-1.55*** (-2.18, -0.92)	-2.43 (-9.45, 4.59)
History_Breathing_Difficulties_Asthma		-0.86*** (-1.20, -0.52)	0.73 (-2.64, 4.10)
History_Cancer	0.12 (-0.30, 0.53)	0.52** (0.08, 0.96)	1.75 (-3.27, 6.77)
History_Diabetes	-2.07*** (-2.40, -1.74)	-1.41*** (-1.77, -1.05)	-1.28 (-4.13, 1.57)
History_Kidney_Condition		-0.77*** (-1.29, -0.26)	-3.68 (-10.71, 3.36)
History_Osteoporosis		-0.78*** (-1.24, -0.32)	-3.30 (-8.25, 1.65)
History_Pacemaker		0.37 (-0.84, 1.58)	-4.42 (-22.99, 14.15)
History_Psychological_Condition		-2.98*** (-3.38, -2.57)	-3.70 (-9.87, 2.47)
History_Seizures		-2.62*** (-3.57, -1.66)	-5.27 (-14.65, 4.11)
History_Smoking	-2.60*** (-2.95, -2.26)	-1.96*** (-2.33, -1.59)	0.62 (-2.74, 3.98)
History_Stroke		-2.78*** (-3.60, -1.95)	-2.02 (-9.41, 5.38)

History_Unexplained_Weight_Loss	-3.30*** (-4.25, -2.35)	-1.89*** (-2.89, -0.89)	2.99 (-6.36, 12.35)
Constant	48.83*** (47.93, 49.73)	52.54*** (51.37, 53.71)	48.97*** (23.35, 74.59)
Summary Statistics			
Observations	82,801	71,662	1,149
R^2	0.40	0.40	0.31
Adjusted R^2	0.40	0.40	0.29
Residual Std. Error	16.50 (df = 82,785)	16.45 (df = 71,636)	18.28 (df = 1,116)
F Statistic	3,670.32*** (df = 15; 82,785)	1,937.78*** (df = 25; 71,636)	15.87*** (df = 33; 1,116)
Note: * $p < .05$, ** $p < .01$, *** $p < .0001$			

Table 4.2: Risk-Adjustment Regression Models for **Neck Patients** by Specification			
	Dependent Variable		
	Outcome Change over Care Episode		
Model Specification	minimal	practical	optimal
Patient-Reported Outcome Change (Median, Interquartile Range)	10 (2, 20)	10 (0, 20)	10 (-1, 21)
Visit Count (Median, Interquartile Range)	12 (8, 17)	12 (8, 17)	14 (9, 19)
	Estimate (95% confidence interval)	Estimate (95% confidence interval)	Estimate (95% confidence interval)
Payer Variables (ref = auto)			
Commercial	-1.46*** (-1.75, -1.17)	-1.15*** (-1.47, -0.84)	0.11 (-2.39, 2.61)
Medicaid	-6.32*** (-6.76, -5.88)	-5.77*** (-6.24, -5.30)	-2.42 (-6.08, 1.23)
Medicare	-0.50** (-0.89, -0.10)	0.21 (-0.21, 0.64)	2.04 (-1.64, 5.71)
Other	-3.52*** (-4.37, -2.67)	-3.03*** (-3.96, -2.10)	-1.10 (-10.24, 8.05)
Workman's Comp	-6.26*** (-6.74, -5.78)	-6.64*** (-7.15, -6.12)	-5.14*** (-8.224, -2.03)
Outcome Score at Index	-0.32*** (-0.33, -0.32)	-0.34*** (-0.35, -0.33)	-0.33*** (-0.39, -0.28)
Health Status at Index			
Height_CM			-0.03 (-0.16, 0.11)
Weight_Lbs			0.001 (-0.02, 0.03)
BMI		0.005 (-0.01, 0.02)	
Chronic (ref = acute)	0.85*** (0.50, 1.20)	-0.36* (-0.76, 0.03)	no variation in data
First PT Visit After Surgery (ref = before)	-1.10*** (-1.55, -0.64)	-1.24*** (-1.72, -0.76)	-6.25*** (-9.89 -2.61)
Pain Score at Index			-0.01 (-1.22, 1.20)
General Mental Score at Index			0.09 (-0.03, 0.22)

General Physical Score at Index			0.06 (-0.09, 0.22)
Socio-demographics			
Male (ref = female)	0.37*** (0.16, 0.59)	0.12 (-0.12, 0.35)	-0.53 (-3.25, 2.19)
Age	-0.07*** (-0.07, -0.06)	-0.05*** (-0.06, -0.04)	-0.16*** (-0.24, -0.09)
Patient Zip Code Median Income	0.07*** (0.02, 0.12)	0.06** (0.01, 0.11)	0.397 (-0.07, 0.82)
African-American (ref = Asian/Pacific)			0.54 (-3.01, 4.08)
Caucasian (ref = Asian/Pacific)			1.59 (-0.59, 3.77)
Eskimo/American Indian (ref = Asian/Pacific)			-0.83 (-7.28, 5.63)
Hispanic (ref = Asian/Pacific)			1.72 (-7.87, 11.31)
Comorbidities at Index			
History_Arthritis		-1.39*** (-1.66, -1.13)	-1.24 (-3.57, 1.09)
History_Blood_Clot		-1.14*** (-1.72, -0.55)	-8.85*** (-15.35, -2.36)
History_Breathing_Difficulties_Asthma		-1.21*** (-1.50, -0.92)	-0.93 (-3.68, 1.82)
History_Cancer	-0.15 (-0.54, 0.25)	0.36* (-0.06, 0.77)	2.11 (-2.33, 6.54)
History_Diabetes	-0.21 (-0.54, 0.11)	0.21 (-0.14, 0.57)	2.90** (0.02, 5.78)
History_Kidney_Condition		-0.75*** (-1.22, -0.28)	-0.40 (-6.10, 5.31)
History_Osteoporosis		-0.44** (-0.85, -0.03)	-2.31 (-6.74, 2.13)
History_Pacemaker		0.14 (-0.93, 1.20)	-0.53 (-14.58, 13.52)
History_Psychological_Condition		-3.04*** (-3.38, -2.70)	-3.76 (-8.79, 1.26)
History_Seizures		-2.22*** (-3.00, -1.45)	-3.81 (-13.54, 5.91)

History_Smoking	-1.61*** (-1.91, -1.31)	-1.08*** (-1.39, -0.76)	-1.29 (-4.31, 1.73)
History_Stroke		-0.48 (-1.20, 0.25)	1.55 (-4.61, 7.71)
History_Unexplained_Weight_Loss	-3.73*** (-4.43, -3.03)	-2.74*** (-3.47, -2.01)	3.41 (-4.67, 11.50)
Constant	35.61*** (34.95, 36.28)	38.05*** (37.10, 38.99)	-37.06*** (-75.54, 62.65)
Summary Statistics			
Observations	66,977	58,164	1,123
R^2	0.16	0.17	0.17
Adjusted R^2	0.16	0.17	0.14
Residual Std. Error	13.46 (df = 66,961)	13.39 (df = 58,138)	15.72 (df = 1.090)
F Statistic	844.75*** (df = 15; 66,961)	473.60*** (df = 25; 58,138)	6.90*** (df = 33; 1,090)
Note: * $p < .05$, ** $p < .01$, *** $p < .0001$			

Table 4.3: Risk-Adjustment Regression Models for **Knee Patients** by Specification			
	Dependent Variable		
	Outcome Change over Care Episode		
Model Specification	minimal	practical	optimal
Patient-Reported Outcome Change (Median, Interquartile Range)	16 (2, 31)	16 (2, 31)	14 (-2, 30)
Visit Count (Median, Interquartile Range)	14 (8, 20)	14 (9, 20)	16 (10, 2)
	Estimate (95% confidence interval)	Estimate (95% confidence interval)	Estimate (95% confidence interval)
Payer Variables (ref = auto)			
Commercial	7.57*** (5.46, 9.67)	7.65*** (5.40, 9.91)	11.28 (-4.25, 26.82)
Medicaid	1.48 (-0.86, 3.83)	2.24* (-0.29, 4.77)	8.73 (-9.93, 26.82)
Medicare	6.38*** (4.17, 8.59)	6.82*** (4.45, 9.18)	0.84 (-16.47, 18.16)
Other	4.66*** (1.68, 7.64)	4.93*** (1.66, 8.21)	17.41 (-10.26, 45.07)
Workman's Comp	1.23 (-1.13, 3.58)	1.24 (-1.30, 3.77)	4.59 (-11.88, 21.33)
Outcome Score at Index	-0.49*** (-0.50, -0.47)	-0.51*** (-0.52, -0.49)	-0.57*** (-0.72, -0.43)
Health Status at Index			
Height_CM			0.01 (-0.46, 0.48)
Weight_Lbs			-0.03 (-0.11, 0.05)
BMI		-0.27*** (-0.32, -0.22)	
Chronic (ref = acute)	-6.08*** (-10.85, -1.31)	no variation in data	no variation in data
First PT Visit After Surgery (ref = before)	5.68*** (4.98, 6.38)	5.21*** (4.44, 5.97)	4.11 (-3.19, 11.41)
Pain Score at Index			2.78 (-1.75, 7.30)
General Mental Score at Index			0.01 (-0.40, 0.43)

General Physical Score at Index			0.01 (-0.50, 0.52)
Socio-demographics			
Male (ref = female)	1.89*** (1.27, 2.52)	1.29*** (0.59, 1.99)	2.23 (-6.97, 11.44)
Age	-0.25*** (-0.27, -0.23)	-0.20*** (-0.22, -0.18)	-0.11 (-0.33, 0.11)
Patient Zip Code Median Income	0.26*** (0.13, 0.39)	0.09 (-0.05, 0.24)	-0.23 (-1.80, 1.35)
African-American (ref = Asian/Pacific)			6.37 (-5.16, 17.90)
Caucasian (ref = Asian/Pacific)			4.11 (-3.11, 11.33)
Eskimo/American Indian (ref = Asian/Pacific)			-4.61 (-30.26, 21.03)
Hispanic (ref = Asian/Pacific)			-3.12 (-27.39, 21.14)
Comorbidities at Index			
History_Arthritis		-1.58*** (-2.39, -0.76)	-1.23 (-10.12 7.65)
History_Blood_Clot		-1.80** (-3.38, -0.23)	19.22* (-2.39, 40.84)
History_Breathing_Difficulties_Asthma		-1.27** (-2.23, -0.30)	2.16 (-7.94, 12.25)
History_Cancer	0.49 (-0.60, 1.58)	0.83 (-0.35, 2.00)	1.92 (-16.63, 20.46)
History_Diabetes	-4.20*** (-5.14, -3.25)	-2.38*** (-3.41, -1.34)	-4.85 (-14.68, 4.97)
History_Kidney_Condition		1.51* (-3.23, 0.20)	6.73 (-20.17, 33.63)
History_Osteoporosis		-2.75*** (-3.95, -1.54)	11.10* (-0.64, 22.84)
History_Pacemaker		-2.15 (-6.07, 1.77)	-19.27 (-63.54, 25.00)
History_Psychological_Condition		-3.21*** (-4.78, -1.63)	-9.08 (-39.24, 21.07)
History_Seizures		-3.85*** (-6.58, -1.13)	-20.77 (-62.80, 21.25)

History_Smoking	-3.59*** (-4.61, -2.56)	-3.38*** (-4.49, -2.28)	-16.32*** (-25.89, -6.75)
History_Stroke		-3.71*** (-5.59, -1.83)	4.95 (-13.48, 23.38)
History_Unexplained_Weight_Loss	-6.19*** (-8.93, -3.44)	-4.52*** (-7.55, -1.49)	-16.41 (-41.28, 8.45)
Constant	51.37*** (45.96, 56.77)	54.19*** (51.07, 57.31)	37.41 (-44.98, 3119.81)
Summary Statistics			
Observations	13,203	10,892	188
R^2	0.34	0.33	0.43
Adjusted R^2	0.34	0.33	0.32
Residual Std. Error	17.71 (df = 13,187)	17.53 (df = 27697)	19.96 (df = 155)
F Statistic	460.22*** (df = 15; 13,187)	553.89*** (df = 25; 27697)	3.70*** (df = 33; 154)
Note: * p < .05, ** p < .01, *** p < .0001			

Table 4.4: Risk-Adjustment Regression Models for **Back Patients** by Specification			
	Dependent Variable		
	Outcome Change over Care Episode		
Model Specification	minimal	practical	optimal
Patient-Reported Outcome Change (Median, Interquartile Range)	8 (-1, 17)	8 (-1, 17)	8 (-1, 17)
Visit Count (Median, Interquartile Range)	12 (8, 17)	12 (8, 17)	13 (9, 18)
	Estimate (95% confidence interval)	Estimate (95% confidence interval)	Estimate (95% confidence interval)
Payer Variables (ref = auto)			
Commercial	0.89*** (0.58, 1.20)	1.48*** (1.15, 1.81)	-0.42 (-2.37, 1.52)
Medicaid	-5.07*** (-5.44, -4.71)	-4.06*** (-4.45, -3.67)	-6.14*** (-8.52, -3.76)
Medicare	0.34* (-0.01, 0.69)	1.30*** (0.93, 1.67)	-0.47 (-2.91, 1.98)
Other	-1.19*** (-1.79, -0.59)	-0.48 (-1.14, 0.17)	-4.11* (-8.62, 0.40)
Workman's Comp	-3.74*** (-4.13, -3.34)	-3.96*** (-4.38, -3.55)	-6.05*** (-8.21, -3.88)
Outcome Score at Index	-0.34*** (-0.34, -0.33)	-0.36*** (-0.36, -0.36)	-0.34*** (-0.38, -0.31)
Health Status at Index			
Height_CM			-0.01 (-0.09, 0.07)
Weight_Lbs			-0.03*** (-0.05, -0.02)
BMI		-0.11*** (-0.12, -0.10)	
Chronic (ref = acute)	0.84*** (0.61, 1.06)	-0.29** (-0.54, 0.03)	no variation in data
First PT Visit After Surgery (ref = before)	-0.06 (-0.34, 0.22)	-0.15 (-0.44, 0.14)	-0.47 (-2.62, 1.68)
Pain Score at Index			-0.27 (-1.03, 0.49)
General Mental Score at Index			0.02

			(-0.06, 0.10)
General Physical Score at Index			0.05 (-0.04, 0.15)
Socio-demographics			
Male (ref = female)	1.25*** (1.11, 1.40)	0.78*** (0.62, 0.93)	1.39* (-0.24, 3.03)
Age	-0.07*** (-0.08, -0.07)	-0.05*** (-0.05, -0.04)	-0.10*** (-0.14, -0.05)
Patient Zip Code Median Income	0.25*** (0.22, 0.28)	0.19*** (0.16, 0.23)	0.27* (-0.01, 0.56)
African-American (ref = Asian/Pacific)			-0.89 (-3.24, 1.47)
Caucasian (ref = Asian/Pacific)			-0.75 (-2.15, 0.65)
Eskimo/American Indian (ref = Asian/Pacific)			-0.35 (-4.42, 3.71)
Hispanic (ref = Asian/Pacific)			-4.80** (-9.44, -0.16)
Comorbidities at Index			
History_Arthritis		-2.40*** (-2.57, -2.22)	-1.94*** (-3.40, -0.48)
History_Blood_Clot		-1.27*** (-1.64, -0.90)	-2.61 (-6.13, 0.91)
History_Breathing_Difficulties_Asthma		-1.41*** (-1.61, -1.21)	-2.98*** (-4.67, -1.28)
History_Cancer	0.11 (-0.14, 0.37)	0.55*** (0.28, 0.82)	2.76* (-0.24, 5.77)
History_Diabetes	-1.76*** (-1.97, -1.56)	-0.74*** (-0.96, -0.52)	0.43 (-1.15, 2.01)
History_Kidney_Condition		-0.76*** (-1.06, -0.47)	0.30 (-2.75, 3.34)
History_Osteoporosis		-0.73*** (-0.98, -0.47)	-1.14 (-3.47, 1.19)
History_Pacemaker		0.15 (-0.48, 0.78)	-6.33 (-14.68, 2.03)
History_Psychological_Condition		-2.97*** (-3.20, -2.74)	-3.31** (-6.00, -0.62)

History_Seizures		-1.50*** (-2.03, -0.96)	-1.42 (-6.09, 3.26)
History_Smoking	-2.66*** (-2.86, -2.46)	-2.00*** (-2.21, -1.79)	-1.53* (-3.24, 0.18)
History_Stroke		-0.67*** (-1.12, -0.22)	1.86 (-1.53, 5.23)
History_Unexplained_Weight_Loss	-2.98*** (-3.46, -2.49)	-2.17*** (-2.68, -1.67)	-2.32 (-5.78, 1.14)
Constant	33.18*** (32.67, 33.70)	39.13*** (38.44, 39.83)	42.34*** (28.13, 56.54)
Summary Statistics			
Observations	152,229	132,215	2,467
R^2	0.15	0.17	0.17
Adjusted R^2	0.15	0.17	0.16
Residual Std. Error	13.85 (df = 152,213)	13.69 (df = 132,189)	14.58 (df = 2,434)
F Statistic	1,804.28*** (df = 15; 152,213)	1,057.54*** (df = 25; 132,189)	15.33*** (df = 32; 2,434)
Note: * p < .05, ** p < .01, *** p < .0001			

Table 5: Partial F-Statistics Indicating the Relative Strength of Baseline Factor Groups in Predicting Outcome Change in the Practical Specification by Body Region.				
	Body Region			
Baseline Factor Group	Shoulder (N = 71,662)	Neck (N = 58,164)	Knee (N = 10,892)	Back (N = 132,215)
Payer Variables[a]	633.6***	294.9***	40.07***	588.8***
Health Status at Index[b]	537.0***	9.8***	Not estimable[e]	121.2***
Socio-Demographics[c]	217.0***	47.0***	109.15***	162.9***
Comorbidities at Index[d]	69.1***	72.7***	16.83***	252.1***

a. Commercial, Medicaid, Medicare, Other, Workman's Comp (reference = auto).
b. BMI, Chronic (ref = acute), First PT Visit After Surgery (ref = before).
c. Male (ref = female), Age, Patient Zip Code Median Income.
d. History_Arthritis, History_Blood_Clot, History_Breathing_Difficulties_Asthma, History_Cancer, History_Diabetes, History_Kidney_Condition, History_Osteoporosis, History_Pacemaker, History_Psychological_Condition, History_Seizures, History_Smoking, History_Stroke, History_Unexplained_Weight_Loss
e. Factor Chronic always equaled 1 for knee patients in sample and statistic could not be estimated.
* $p < .05$, ** $p < .01$, *** $p < .0001$

Figure 2: Physical Therapist Performance Measures for Shoulder Patients (physical therapists treating a minimum of 40 shoulder patients)
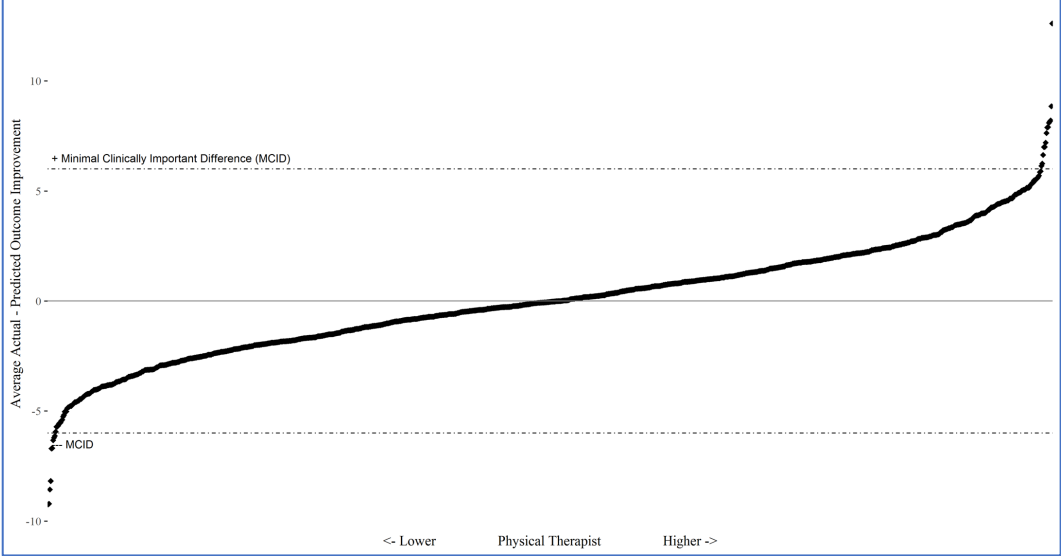
Figure 3: Physical Therapist Performance Measures for Neck Patients (physical therapists treating a minimum of 40 neck patients)
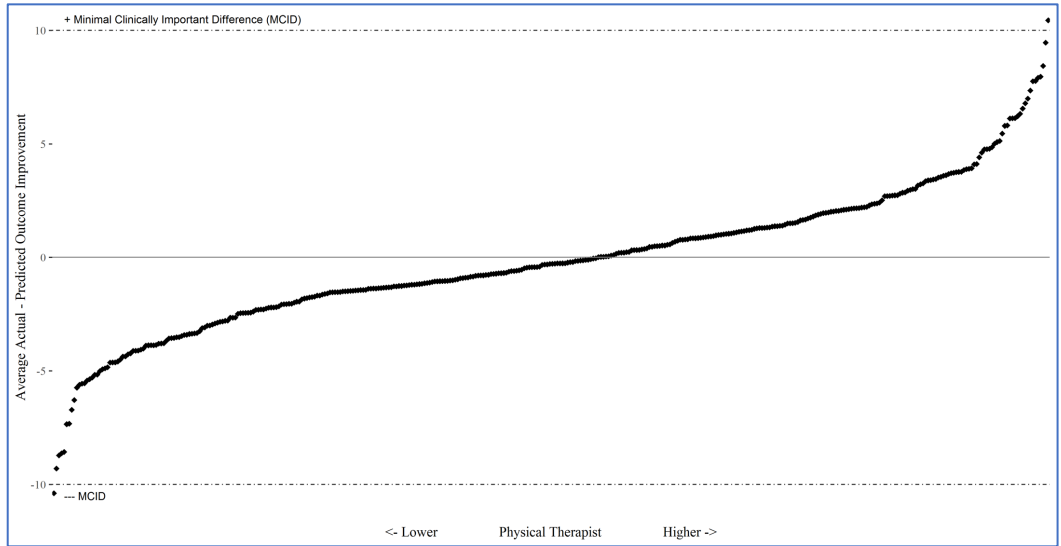
Figure 4: Physical Therapist Performance Measures for Knee Patients (physical therapists treating a minimum of 40 knee patients)
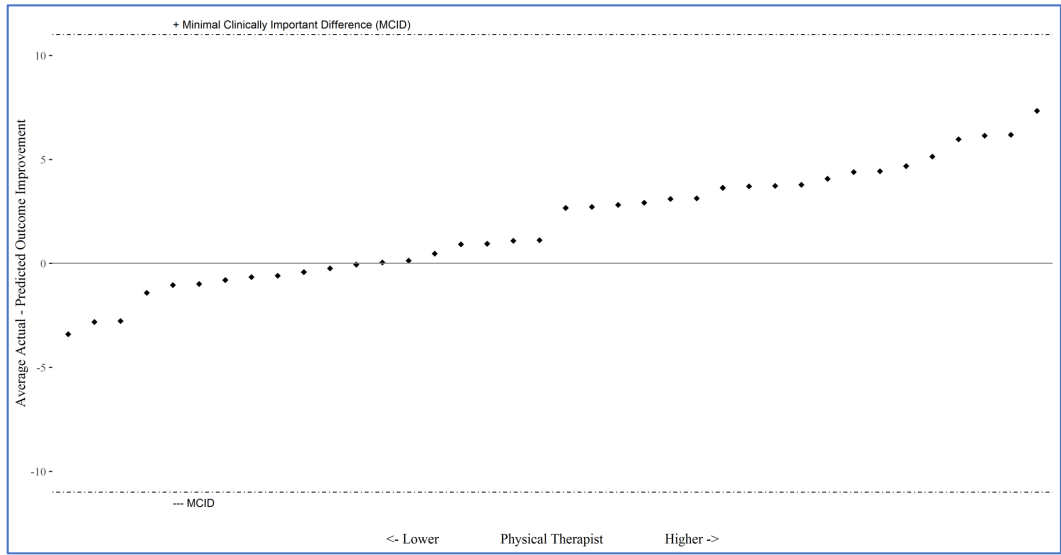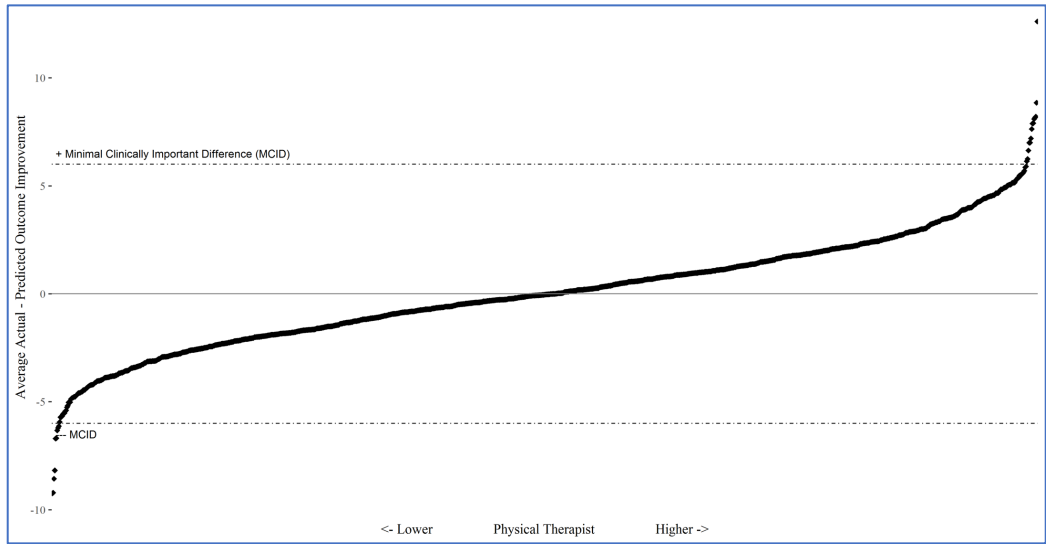
Figure 5: Physical Therapist Performance Measures for Back Patients (physical therapists treating a minimum of 40 back patients)

References

1.   Gummesson C, Ward MM, Atroshi I. The shortened disabilities of the arm, shoulder and hand questionnaire (QuickDASH): validity and reliability based on responses within the full-length DASH. *BMC Musculoskelet Disord.* 2006;7:44.
2.   Binkley JM, Stratford PW, Lott SA, Riddle DL. The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. North American Orthopaedic Rehabilitation Research Network. *Phys Ther.* 1999;79(4):371-383.
3.   Fritz JM, Irrgang JJ. A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. *Phys Ther.* 2001;81(2):776-788.
4.   Stratford PW, Binkley JM. Applying the results of self-report measures to individual patients: an example using the Roland-Morris Questionnaire. *The Journal of orthopaedic and sports physical therapy.* 1999;29(4):232-239.
5.   Stratford PW, E F, Soloman P, Binkley JM, Gill C, Moreland J. Using the neck disability index to make decisions about individual patients. *Physiother Can.* 1996;48:107-110.
6.   Polson K, Reid D, McNair PJ, Larmer P. Responsiveness, minimal importance difference and minimal detectable change scores of the shortened disability arm shoulder hand (QuickDASH) questionnaire. *Man Ther.* 2010;15(4):404-407.
7.   U.S. Census Bureau. American Community Survery, 2016 American Community Survey 5-Year Estimates, Table S1901. 2018. Accessed March 1, 2018.