

Predicting Risk of Adverse Outcomes in Knee Replacement Surgery with Reconstructability Analysis

Cecily Froemke, PhD
Systems Science Program, Portland State University;
Providence Health & Services
Portland OR, USA
cfroemke@gmail.com

Martin Zwick, PhD
Systems Science Program,
Portland State University
Portland OR, USA
zwick@pdx.edu

Abstract—Reconstructability Analysis (RA) is a data mining method that searches for relations in data, especially non-linear and higher order relations. This study shows that RA can provide useful predictions of complications in knee replacement surgery.

Keywords—Reconstructability Analysis, Occam, predictive analytics, healthcare, risk prediction, total knee replacement.

I. INTRODUCTION

Legislative reforms aimed at slowing growth of US healthcare costs are focused on achieving greater value per dollar. To increase value while payments are diminishing and tied to individual outcomes, healthcare providers must do better at predicting risks and outcomes. One way to improve predictions is through enhanced modeling methods. Current modeling is predominantly done with logistic regression (LR). This project applied Reconstructability Analysis (RA) to data on hip and knee replacement surgery to predict complications in patient outcomes, and this paper reports a few of the results of the knee study. RA is partially similar to LR, but has some unique features.

RA is a data mining method that searches for relations in data, especially non-linear and higher ordinality relations, by decomposing the frequency distribution of the data into projections, several of which taken together define a model, which is then assessed for statistical significance. The predictive power of the model is expressed as the percent reduction of uncertainty (Shannon entropy) of the dependent variable (the DV) gained by knowing the values of the predictive independent variables (the IVs). Here we report the prediction of complications (DV), given a set of patient comorbidities (IVs). Prediction is done with the conditional probability distribution of the DV given the IVs specified by an RA model of the data. Complex interaction effects between the IVs and the DV may allow better predictions than predictive IVs used separately. Exploratory modeling with RA may even detect novel and surprising predictors. The main virtue of exploratory modeling is that relations between the IVs and the DV do not have to be specified up front, and thus their form does not need to be known or hypothesized. Relations can be discovered. For example, in a study applying RA to genomic data, researchers found that RA can detect gene-gene interactions that other methods could not detect [1].

II. METHODS

A. Reconstructability Analysis

RA developed from the early works of Ross Ashby [1] who defined a process for systematically testing whether a complex constraint could first be decomposed into several simpler constraints and then, using the maximum entropy principle, recomposed without suffering serious information loss. RA assesses the goodness of models that are hypergraphs either using set theoretic (SRA) or information theoretic (IRA) measures. IRA, the approach used in this project, resembles log-linear statistical methods in the social sciences, and has had diverse applications including time-series analysis, classification, decomposition, compression, pattern recognition, prediction, control, and decision analysis [3]. Several RA software applications exist such as GSPS [4], Construct and Spectral [5], SAPS [6], EDA [7] and Occam [8 9]. For this project, the Occam software was used.

Although it is designed for nominal multivariate data, RA can also handle continuous data by binning values into discrete binary or multi-valued states. The more states of an IV the better it can predict the outcome, but as the number of states of a variable increases the sample size required also increases, so the number of bins used for variables is a scarce resource that must be allotted judiciously.

To illustrate the IRA method, consider data on four variables, three IVs (A, B, C) and one DV (Z). For these four variables, multiple relations are possible, and each set of non-redundant relations is a graph or hypergraph structure that is a candidate model of the data. There are 19 such structures for three IVs and one DV, and for such a small number of variables, exhaustive search of all models is possible. In the current project, there are 188 IVs, which generate a massive lattice of structures which cannot be examined exhaustively but must instead be searched with intelligent heuristics.

Search for predictive models that are statistically significant begins with the independence model, which for our illustrative example is $ABC:Z$. This model says that there may or may not be a relation among the IVs (A, B, C), but none of the IVs predict Z. An ascending search then examines increasingly complex – and more predictive – models until difference from independence and gains in uncertainty reduction due to increases of complexity are no longer statistically significant. For example, one possible model that

the search might yield is ABC:ABZ:CZ; this model contains an ABZ component that represents a predictive interaction effect of two IVs, A and B, and the DV, plus an additional predictive relation of C with the DV. Model search is done at two levels of refinement: variable-based model without loops (a “coarse” search) and variable-based models with loops (a “fine” search), the refined search yielding more predictive and typically more complex models.

To avoid overfitting, i.e., choosing an overly complex model that does poorly when confronted with new data, a good model should capture maximum information (constraint) in the data while being as simple as possible. A simple model is one whose degrees of freedom are not much greater than the independence model. In Occam, the tradeoff between information-captured and simplicity is done using three different criteria: the Bayesian Information Criterion (BIC), the Akaike Information Criterion (AIC), and the Incremental-p Chi-square criterion (IncrP). BIC and AIC aggregate information-captured and simplicity linearly, with BIC penalizing models for complexity more than AIC. The third criterion, IncrP, selects the model with the highest reduction of DV uncertainty, where the difference between the model and independence is statistically significant and where, in addition, there is a path from independence where each incremental step to the model is also significant. (A p-value of 0.05 was used as the cutoff for significance.) IncrP is sometimes more conservative than AIC, sometimes less conservative, but BIC is always the most conservative of the three, and in this study, was the criterion used to select the “best” model. BIC is reported below in TABLE 1 as the difference between BIC for independence and BIC for the model. The table also reports the percent reduction of uncertainty of the DV achieved by the model, $\% \Delta H(DV)$, which is the actual predictive power of the model. Calculation of uncertainty does not involve the sample size and is non-statistical [10]; its significance is assessed by its p-value or by the BIC/AIC measure. The reduction in uncertainty, a central measure of RA not generally available with other methods, is more sensitive to the predictive strength of a model than $\% \text{correct}$ and related measures. Because of the logarithm term in the expression for uncertainty, even small reductions of uncertainty can correspond to big effect sizes. For example, an 8% reduction of uncertainty can correspond to a shift in the odds of possible outcomes as big as a change from 1:1 to 2:1.

After the best model is obtained, its actual contents – what predictions it makes for the DV for all the different IV states – is examined in detail. In Occam, this detailed examination is called “fit,” to be distinguished from the first step which is called “search.” Search results below are shown in TABLE 1, fit results in TABLE 2. For more information about RA, see [3] and [11]. For more information about this study, see [12], which also includes a demonstration that RA provides predictive results not available from logistic regression.

B. The Data

Data used in this study derives from patients who underwent an inpatient surgical procedure of a total knee replacement at one of seven inpatient hospitals within an

integrated healthcare system in a single state. Participant data consists of both hospital billing data and electronic health record system clinical data. Clinical and cost data were matched on the patient’s episode identifier, then de-identified and transformed into the variables used in this research project. Because the administrative claims database includes variables that are collected in diverse health systems across the nation, the resulting predictive model developed in this project have the potential for wide-spread use.

There are 4,336 cases in the knee data set. ICD-9 codes were used to classify the procedure of an elective total knee replacement procedure (81.54) and to classify the comorbidity IVs and the DV Complication occurring for each knee procedure. The independent variables age (Age), surgeon volume (Sv), and number of risks (Nr) were continuous variables that were discretized into the binned variables Ageb, Svb, and Nrb. These IVs were divided into 3 bins, with equal sample sizes to allow optimal predictive capacity. The DV Complication (Cp) was created by looking at the ICD-9 diagnosis codes with a Present On Arrival indicator of 0, indicating the diagnosis was acquired after admission to the hospital. The knee data set contained 913 complications in 205 cases. The complication rate for the knee data set is thus 205/4336 or 4.7%.

Preliminary analyses indicated the need to reduce the set of IVs. This was done with a level = 1 loopless search which assessed the predictive strengths, expressed in $\% \Delta H$ reduction, of the 188 IVs. An IV was retained if its p value was $\leq .05$. Sorting IVs by $\% \Delta H$ showed the single IVs with the greatest predictive strength.

Initially analyses were conducted with training/ test splits, but these resulted in $\% \text{correct}$ measures that were small and misleading. While training/test splits is common in machine learning research, it is often done with larger sample sizes and fewer variables. This project’s primary objective was exploratory modeling, whose results need to be subjected to subsequent confirmatory testing. Training/test splits were thus not considered to be necessary.

III. RESULTS

A. Model Search

A model in TABLE 1 specifies the IVs (e.g., Nrb, Rku) that predict the DV (Cp), followed by $\Delta df = df(\text{model}) - df(\text{reference})$, the difference in degrees of freedom of the model and independence; then $\Delta BIC = BIC(\text{reference}) - BIC(\text{model})$, for which improvements in the model compared to the reference are reflected in larger positive values; then $\% \Delta H = 100 (H(DV) - H(DV|IV)) / H(DV)$, the $\% \text{reduction}$ of uncertainty of the DV given the IVs. The reduction of uncertainty measure indicates how predictive the IVs are, while the BIC measure indicates how *efficient* the prediction is, i.e., how predictive the IVs are, given their complexity (df). Best models are chosen based on their ΔBIC values, which results in a highly conservative model choice.

TABLE 1 summarizes the results of single and multiple predictors in loopless and all-model (with loops) searches. The best coarse model shows that, for this data set, simply knowing the *total number of comorbidities* a patient had (Nrb) along

with *chronic kidney disease* (Rku) reduces the uncertainty in predicting if Complication (**Cp**) occurred by 7.58%. Knowing the surgeon who performed the surgery (S) reduces uncertainty by 6.45%. Likewise, knowing only if the patient had *unspecified hypertensive renal disease* (Rrd) reduces uncertainty by 3.11%.

The next type of search considers models with loops which allows for multiple components predicting the DV. Within each component, there may be interaction effects among the IVs in their prediction of the DV, just as interaction effects were observed in the best loopless BIC and AIC/IncrP models, Nrb Rku **Cp** and Nrb Rhd Rku **Cp**, shown in TABLE 1.

Note that some *single* predicting variables do not show up in the best coarse or fine models, indicating that the IVs are not independent from each other. There are 6 single predicting variables in the best BIC fine-grained model, Ageb **Cp** : Nrb **Cp** : Ruh **Cp** : Rhd **Cp** : Rku **Cp** : Rro **Cp**. Five of these variables – Ageb, Nrb, Ruh, Rhd, and Rku – also appear in the

top 10 single predicting components, while Rro is the 18th in the list of single predicting components. This apparently low-value variable was included when the RA search methodology sought to improve a model already containing the better individual predictors Ageb, Nrb, Ruh, Rhd, and Rku. Rro was found to be the variable that added more *additional* information to that model than any of the better single-predicting IVs above it.

The best single predictor, S (surgeon) does not appear in the best fine-grained model, presumably in part because S has high cardinality and the information added by S is not worth the complexity of including it in the model and perhaps in part also because the predictive effect of S is already provided by the Ageb, Nrb, Ruh, Rhd, and/or Rku predictors. Similarly, Ageb, Nrb, Ruh, Rhd, and Rku contain the information offered by the other single predictors all the way down to Rro.

TABLE 1. Summary of Search Results for All IVs. Search covers coarse and fine models. All p-values = 0.

MODEL	Δdf	ΔBIC	%ΔH	Variable description
COARSE, single predictors (top 10)				
S Cp	62	-412.7	6.45	Surgeon
Nrb Cp	2	77.29	5.69	Number of risks (binned)
Rrd Cp	1	43.04	3.11	Unspecified hypertensive renal disease (403.9)
Rku Cp	1	39.63	2.91	Chronic kidney disease, unspecified (585.9)
Ruh Cp	1	33.56	2.54	Other and unspecified hyperlipidemia (272.4)
L Cp	6	-9.04	2.5	Location
Ad Cp	27	-185.3	2.47	Admission diagnosis
Ageb Cp	2	14.61	1.9	Age (binned)
Raf Cp	1	11.46	1.2	Atrial fibrillation (427.31)
Rhf Cp	1	10.79	1.16	Heart failure (428)
MODEL	Δdf	ΔBIC	%ΔH	Variable description
COARSE, single predictors not in the top 10 but in AIC or BIC models below				
Rhd Cp (rank 12)	1	9.9	1.11	Other chronic pulmonary heart disease (416.8)
Rro Cp (rank 18)	1	3.22	0.7	Rosacea (695.3)
Reg Cp (rank 20)	1	1.95	0.63	Esophagitis (530.1)
MODEL	Δdf	ΔBIC	%ΔH	Variable description
COARSE, best model (loopless)				
ΔBIC (best model)				
Nrb Rku Cp	5	83.23	7.58	Number of risks (binned), Chronic kidney disease, unspecified (585.9)
IncrP & ΔAIC (same best model)				
Nrb Rhd Rku Cp	11	52.71	8.77	Number of risks (binned), Other chronic pulmonary heart disease (416.8), Chronic kidney disease (585.9)
MODEL	Δdf	ΔBIC	%ΔH	Variable description
FINE, best models (with loops)				
ΔBIC (best model)				
Ageb Cp : Nrb Cp : Ruh Cp : Rhd Cp : Rku Cp : Rro Cp	8	104.7	10.4	Age (binned), Number of risks (binned), Other and unspecified hyperlipidemia (272.4), Other chronic pulmonary heart disease (416.8), Chronic kidney disease, unspecified (585.9), Rosacea (695.3)
IncrP & ΔAIC (same best model)				
Ageb Cp : Nrb Cp : Ruh Cp : Rhd Cp : Reg Cp : Rku Cp : Rro Cp	9	104.2	10.88	Age (binned), Number of risks (binned), Other and unspecified hyperlipidemia (272.4), Other chronic pulmonary heart disease (416.8), Esophagitis (530.1), Chronic kidney disease, unspecified (585.9), Rosacea (695.3)

The third best single predictor, Rrd, does not appear in the best fine-grained model either. Again, the information it would add is presumably not worth the additional complexity it would add. This explanation is supported by the fact that Rrd is well predicted by Ageb, Nrb, Ruh, Rhd, and Rku. In fact, Rku alone predicts Rrd with a $\% \Delta H$ of 53.14% demonstrating significant overlap between Rku and Rrd. This lack of independence between the IVs is analogous to collinearity among IVs in regression analysis.

The next type of search considers models with loops which allows for multiple components that predict the DV. Unlike the best *loopless* models shown in TABLE 1, the best model for **Cp** now does not contain interaction terms.

B. Model Fit

Having found a best model, the next step is to analyze its detailed content- i.e., the conditional probability distribution for the DV, given the predicting IVs. This distribution is shown in TABLE 2 for the best fine-grained model Ageb **Cp** : Nrb **Cp** : Ruh **Cp** : Rhd **Cp** : Rku **Cp** : Rro **Cp**.

TABLE 2 Fit Table for Best Model: Ageb **Cp** : Nrb **Cp** : Ruh **Cp** : Rhd **Cp** : Rku **Cp** : Rro **Cp**. Blue rows are for ratio < 0.90, orange rows for ratio > 1.10.

IVs							Data			Model			
							<i>obs. p(DV IV)</i>			<i>calc. q(DV IV)</i>			
#	Ageb	Nrb	Ruh	Rhd	Rku	Rro	freq	Cp=0	Cp=1	Cp=0	Cp=1	ratio	p(margin)
1	1	1	0	0	0	0	502	99.00	1.00	99.11	0.89	0.19	0.00
2	1	1	1	0	0	0	1	100.00	0.00	98.46	1.54	0.33	0.88
3	1	2	0	0	0	0	457	98.69	1.31	97.77	2.24	0.47	0.01
4	1	2	0	0	0	1	1	100.00	0.00	80.86	19.14	4.05	0.50
5	1	2	0	0	1	0	2	100.00	0.00	91.86	8.14	1.72	0.82
6	1	2	0	1	0	0	1	0.00	100.00	87.38	12.62	2.67	0.71
7	1	2	1	0	0	0	34	91.18	8.82	96.17	3.83	0.81	0.81
8	1	3	0	0	0	0	380	96.05	3.95	95.90	4.10	0.87	0.56
9	1	3	0	0	0	1	1	100.00	0.00	69.34	30.66	6.48	0.22
10	1	3	0	0	1	0	8	100.00	0.00	85.80	14.20	3.00	0.24
11	1	3	0	1	0	0	2	100.00	0.00	78.75	21.25	4.49	0.27
12	1	3	1	0	0	0	96	89.58	10.42	93.07	6.93	1.47	0.31
13	1	3	1	0	0	1	1	100.00	0.00	56.47	43.53	9.21	0.07
14	1	3	1	0	1	0	3	66.67	33.33	77.61	22.39	4.74	0.15
15	1	3	1	1	0	0	1	0.00	100.00	68.01	31.99	6.77	0.20
16	2	1	0	0	0	0	421	99.29	0.71	98.78	1.22	0.26	0.00
17	2	1	0	1	0	0	1	100.00	0.00	92.78	7.22	1.53	0.91
18	2	1	1	0	0	0	6	100.00	0.00	97.90	2.10	0.44	0.76
19	2	2	0	0	0	0	420	96.91	3.10	96.96	3.04	0.64	0.10
20	2	2	1	0	0	0	50	90.00	10.00	94.82	5.18	1.10	0.88
21	2	3	0	0	0	0	349	93.98	6.02	94.47	5.53	1.17	0.48
22	2	3	0	0	0	1	3	33.33	66.67	62.26	37.74	7.98	0.01
23	2	3	0	0	1	0	10	60.00	40.00	81.51	18.49	3.91	0.04
24	2	3	0	1	0	0	3	66.67	33.33	73.00	27.00	5.71	0.07
25	2	3	0	1	1	0	1	100.00	0.00	41.10	58.90	12.46	0.01
26	2	3	1	0	0	0	137	95.62	4.38	90.74	9.26	1.96	0.01
27	2	3	1	0	1	0	9	44.44	55.56	71.66	28.34	5.99	0.00
28	2	3	1	1	0	0	1	100.00	0.00	60.80	39.21	8.29	0.11
29	3	1	0	0	0	0	376	97.87	2.13	98.11	1.90	0.40	0.01
30	3	1	1	0	0	0	2	50.00	50.00	96.74	3.26	0.69	0.92
31	3	2	0	0	0	0	447	95.08	4.92	95.32	4.68	0.99	0.96
32	3	2	0	0	0	1	1	0.00	100.00	66.30	33.71	7.13	0.17
33	3	2	0	0	1	0	7	100.00	0.00	84.01	15.99	3.38	0.19
34	3	2	1	0	0	0	54	94.44	5.56	92.11	7.89	1.67	0.27
35	3	3	0	0	0	0	341	90.62	9.38	91.60	8.41	1.78	0.00
36	3	3	0	0	0	1	2	100.00	0.00	51.29	48.71	10.30	0.00
37	3	3	0	0	1	0	28	75.00	25.00	73.77	26.23	5.55	0.00
38	3	3	0	1	0	0	7	57.14	42.86	63.31	36.70	7.76	0.00
39	3	3	0	1	1	0	1	100.00	0.00	30.81	69.19	14.63	0.00
40	3	3	1	0	0	0	148	87.84	12.16	86.21	13.79	2.92	0.00
41	3	3	1	0	0	1	1	0.00	100.00	37.65	62.35	13.19	0.01
42	3	3	1	0	1	0	18	66.67	33.33	61.74	38.26	8.09	0.00
43	3	3	1	1	0	0	2	50.00	50.00	49.74	50.26	10.63	0.00
							4336	95.27	4.73	95.27	4.73	1.00	
#	Ageb	Nrb	Ruh	Rhd	Rku	Rro	freq	Cp=0	Cp=1	Cp=0	Cp=1	ratio	p(margin)

The columns of the table are: the model number, to be able to refer to models easily; the six IVs in the model and their different states; the frequency of each particular IV (vector) state; the conditional probability $p(\mathbf{Cp}=0|\mathbf{IV})$ and $p(\mathbf{Cp}=1|\mathbf{IV})$ in the data given as percentages; these two conditional probabilities in the *model*, written as $q(\mathbf{Cp}=0|\mathbf{IV})$ and $q(\mathbf{Cp}=1|\mathbf{IV})$; the ‘risk ratio’ of $q(\mathbf{Cp}=1|\mathbf{IV}) / q(\mathbf{Cp}=1)$, i.e., the probability of complications for a particular IV state divided by the marginal probability of complications for the whole sample. So, for example, the first row specifies the IV state (Ageb, Nrb, Ruh, Rhd, Rku, Rro) = (1,1,0,0,0,0), which occurs 502 times in the sample, for which the conditional probabilities for the data (p) and the model (q) are given in percent, where risk ratio $0.19 = 0.89/4.73$, and where the p-value for the comparison of (99.11, 0.89) to the margins (95.27, 4.73) is 0. The ‘risk ratio’ conveys the *effect size*, while the p-value conveys the *significance* of the effect size.

For the independence model, which is the reference, we do not know the state of Ageb or Nrb or if a comorbidity was present, so the uncertainty of the DV comes from its marginal distribution, which is the last line of the table, for which the data and model conditional probabilities are the same. For the calculated model, knowing the states of Nrb and Ageb and the presence or absence of individual comorbidity IVs (Ruh, Rhd, Rku, Rro) tells us about the probability of a complication occurring. Model conditional probabilities are more appropriate to use than data conditional probabilities because the model is simpler than the data and generalizes better.

The marginal distribution (last line) of TABLE 2 shows that in the sample of 4,336 knee replacement cases, Complication ($\mathbf{Cp}=1$) was present in 4.73% and absent in 95.27% of the cases. If the conditional probabilities for particular IV states are either higher or lower than the margins, then the IVs have provided new (predictive) information. Looking at TABLE 2 shows a number of rows whose calculated probabilities are very different from the margins: the blue and orange shaded cells. Rows are highlighted if $p(\text{margin}) \leq 0.05$ and frequency >10 . Aside from very low-frequency IV states (rows 25, 39, and 41), the model distribution never predicts more than a 50% chance of $\mathbf{Cp} = 1$, i.e., it always predicts $\mathbf{Cp} = 0$, which is just what the marginal distribution predicts even without any IV information. The additional information that the model provides beyond the independence model is the *risk* of complication occurrence. While there were no IV states with sizeable frequencies where $q(\mathbf{Cp}=1|\mathbf{IV}) > 0.5$, there are probabilities that are considerably different than the margins, which demonstrate a lower ($< 4.73\%$) or higher ($> 4.73\%$) risk of complications. These deviations from the risk of the overall sample are indicated by the risk ratio: when ratio is < 0.90 (and statistically significant), risk is reduced (blue cells), compared to the margins; when ratio > 1.10 (and statistically significant), risk is increased (orange cells).

Row 1, for example, shows a protective effect for age < 63 (bin=1 for *age binned*, Ageb) and number of risks ≤ 1 (bin = 1 for *number of risks binned*, Nrb) where the probability of $\mathbf{Cp}=1$ is 0.89% (ratio = 0.19), markedly lower than the margin of 4.73%. Row 16 shows a similar protective effect, where

even with age range 63-71 (bin = 2 for Ageb), as long as the number of risks ≤ 1 (bin = 1 for Nrb), the probability is 1.22%, which is lower than the margin (ratio = 0.26). Row 29 also offers a protective effect where even with age range 72-95 (bin = 3 for Ageb) as long as the number of risks ≤ 1 (bin = 1 for Nrb) then the probability of $\mathbf{Cp}=1$ is still lower than the margin at 2.13% (ratio = 0.40). Row 3 shows that even where there is an increase in number of comorbidities with number of risks = 2 or 3 (bin = 2 for Nrb), when Ageb=1, there is still a protective effect with probability of $\mathbf{Cp}=1$ of 2.24% (ratio = 0.47). In each of these three cases where there was a protective effect, the four comorbidity IVs, Ruh, Rhd, Rku and Rro were all absent. To recapitulate: the results show that if these comorbidity IVs are absent and Nrb = 1, then Ageb can be in any of its 3 potential states and the risk is still low. Risk is also reduced if Ruh, Rhd, Rku and Rro are not present, even if there are more comorbidities present (Nrb = 2) if the age is low (Ageb = 1).

Row 35 shows IV states that predict *higher* risk of $\mathbf{Cp}=1$. With age range 72-95 (bin = 3 for Ageb), and number of risks between 4 and 18 (bin = 3 for Nrb), there is a higher probability of $\mathbf{Cp}=1$, namely 8.41% (ratio = 1.78). In this state, there was no presence of one of the four comorbidity IVs (Ruh, Rhd, Rku & Rro). In row 23, however, with the presence of Rku and with lower age 63-71 (bin = 2 for Ageb), and with number of risks between 4 and 18 (bin = 3 for Nrb), the probability of $\mathbf{Cp}=1$ is 18.49% (ratio = 3.91). Compare row 35 also with row 37 in TABLE 2 (freq = 28) where again, Ageb = 3 and Nrb = 3 but Rku is present and we get a much higher risk ratio of 5.5, a 0.2623 probability of $\mathbf{Cp}=1$ which is over 5 times the risk of the whole sample.

A complication ($\mathbf{Cp}=1$) was observed in 4.73% (205 patients) of the 4336 patients in the knee data set, so this is the percentage of patients for which the independence model, which takes into account nothing about the patients or the healthcare delivery system, would thus predict complication. However, the best model from this analysis (Ageb \mathbf{Cp} : Nrb \mathbf{Cp} : Ruh \mathbf{Cp} : Rhd \mathbf{Cp} : Rku \mathbf{Cp} : Rro \mathbf{Cp}) identified several groups of patients who were at increased risk of \mathbf{Cp} with particular combinations of IV states from the model. Considering these high-risk groups together, 15.73% of the total patients in the sample had an increased risk of complication. For these patients at increased risk, the weighted average risk ratio is 2.41; thus 11.40% (or 494 patients) out of that group (15.73% of the whole sample) would be predicted to experience a complication.

IV. DISCUSSION & CONCLUSIONS

Predictive models can augment clinical decision making by providing additional information. The models resulting from this research provide new information about risk for a sizeable proportion of the patient population. If used in real time, such risk predictions could support clinical decision making and custom tailored utilization of services.

One of the purposes of this research project was to determine the variables that were the most predictive of each of the DVs. A sample of previously known-to-be-predictive

IVs were included in the data sets for this project; results validated many of these as important predictors while excluding others. Additionally, the exploratory modeling approach used in this project sought to detect novel or surprising IVs that may not have been hypothesized previously in the literature. Indeed, a number of novel IVs were found to be important.

Future research might rectify the limitations of this project's data and employ additional RA techniques and training-test splits. Implementation of predictive models should be discussed with considerations for data supply lines, maintenance of models, organizational buy-in, and the acceptance of model output by clinical teams for use in real time clinical practice.

This project demonstrated that RA can be useful in the prediction of complications for knee replacement surgery. It also has implication for broader testing and applications. RA is likely to be useful for constructing predictive models for other outcomes of interest and in other clinical areas.

If outcomes and risk are adequately predicted, areas for potential improvement become clearer, and focused changes can improve patient care. Better predictions, such as those resulting from the RA methodology, can thus support improvement in healthcare value – better outcomes at a lower cost.

As reimbursement increasingly evolves into value-based programs, understanding the outcomes achieved, and customizing patient care to reduce unnecessary costs while improving outcomes, will be an active area for clinicians, healthcare administrators, researchers, and data scientists for years to come.

ACKNOWLEDGMENT

We thank Dr. Joe Fusion for edits, suggestions, and endless encouragement.

REFERENCES

- [1] S. Shervais, P. L. Kramer, S. K. Westaway, N.J. Cox., & M. Zwick, Reconstructability analysis as a tool for identifying gene-gene interactions in studies of human diseases,” *Statistical Applications in Genetics & Molecular Biology*, 9(1), 2010, pp. 1–25.
<https://www.pdx.edu/sites/www.pdx.edu/sysc/files/SAGMB.pdf>
- [2] R. Ashby, “Constraint analysis of many-dimensional relations,” *General Systems Yearbook*, 9, 1964, pp. 99–105.
- [3] M. Zwick, “An overview of reconstructability analysis,” *Kybernetes: The International Journal of Systems & Cybernetics*, 33(5/6), 2004, pp. 877–905.
<https://www.pdx.edu/sysc/sites/www.pdx.edu/sysc/files/overview.pdf>
- [4] G. Klir, *The Architecture of Systems Problem Solving*. New York: Plenum Press, 1985.
- [5] K. Krippendorff, “An algorithm for identifying structural models of multivariate data,” *International Journal of General Systems*, 7(1), 1981, pp. 63–79.
- [6] F. E. Cellier & D. W. Yandell, “SAPS-II: A new implementation of the Systems Approach Problem Solver,” *International Journal of General Systems International Journal of General Systems*, 13(4), 1987, pp. 307–322.
- [7] R. C. Conant, “Extended dependency analysis of large systems,” *International Journal of General Systems International Journal of General Systems*, 14(2), 1988, pp. 97–123.
- [8] K. Willett & M. Zwick, “A software architecture for reconstructability analysis,” *Kybernetes: The International Journal of Systems & Cybernetics*, 33(5/6), 2004, pp. 997–1008.
- [9] M. Zwick, “OCCAM: a reconstructability analysis program,” 2016.
<https://www.pdx.edu/sysc/sites/www.pdx.edu/sysc/files/woccaman-2.27.2016.pdf>
https://www.pdx.edu/sites/www.pdx.edu/sysc/files/sysc_kenpitf.pdf
- [10] M. Zwick, “Reconstructability analysis of epistasis,” *Annals of Human Genetics*, 75(1), 2011, pp. 157–171. <http://doi.org/10.1111/j.1469-1809.2010.00628.x>
https://www.pdx.edu/sites/www.pdx.edu/sysc/files/AHG_final_unformatted-I.pdf
- [11] <https://www.pdx.edu/sysc/research-discrete-multivariate-modeling>
- [12] C. C. Froemke, PhD dissertation: Enhancing Value-Based Healthcare with Reconstructability Analysis: Predicting Risk for Hip and Knee Replacements, Portland State University, 2017.
http://pdxscholar.library.pdx.edu/open_access_etds/3772/