Compassion Fatigue and its Relationship to Coping: Statistical Analysis of a Trauma Center Trauma Sensitive Yoga (TCTSY) Study

Matt Parker

School of Statistics, University of Minnesota

STAT 4893W: Consultation and Communication for Statisticians

Professor Kuzmak

October 12, 2023

Abstract

Using data provided by Dr. Nguyen-Feng, our statistical research will focus on coping strategies and compassion fatigue of TCTSY'F's. Our objective is to determine whether a statistically significant relationship exists between compassion fatigue of TCTSY-F's and the coping strategies they employ. Our study implemented two linear regression models for each compassion fatigue variable (job burnout and secondary trauma) while using coping strategies as covariates. We converted our coefficients in the model to beta coefficients, which are commonly used in psychology. We implemented bagging regression trees as another form of analysis to determine the importance of each coping strategy in the models. Ultimately, we were not able to make statistically significant conclusions on our regression model, as beta coefficients are outside of the realm of statistical analysis, but we were able to make distinctions between our covariates using cluster analysis, as well as our observations using k-means clustering.

Introduction

Through a comprehensive and detail driven presentation, STAT 4893W had the pleasure of learning about Trauma-Sensitive Yoga Facilitators (TCTSY-F's) and their impact in the trauma care community. We learned about the safe and empowering environments they create, along with the clinically based solutions they provide to trauma survivors. The data collected on TCTSY-F's highlighted the mental strain involved in trauma care; However, we can go further by applying statistical analysis to the relationship between coping and fatigue.

Our research objective is to determine whether specific coping strategies have a statistically significant impact in predicting compassion fatigue. Oftentimes, statistical analysis highlights that only a select few of our covariates (coping strategies) in our model are significant in predicting our response. By conducting this analysis, we will be able to highlight the most and least significant coping strategies for TCTSY-F's, which is imperative for the study, as these results may be a stepping stone towards furthering care for trauma care providers.

Furthermore, our statistical analysis will address two key questions: How would we rank the coping strategies of TCTSY-F's in relation to compassion fatigue, and how can we distinguish between adaptive and maladaptive coping strategies?

Methods and Materials

Our statistical analysis will employ data collected from an observational study. Given that the 60 coping strategy items were split into 15 groups of 4, with each group being associated with a specific strategy, we used the sum of each group score as our covariates.

The data set provides two response variables: Job burnout and secondary trauma. We use the fifteen covariates (coping strategies) and regress these covariates onto our response. Our responses will be the log-transformed burnout and secondary trauma scores.

We will conduct this analysis using two multiple linear regression models. MLR's are used to evaluate the relationship between two or more predictor variables and one response variable. While it's typically important to highlight the assumptions and hypotheses associated with MLR's, we cannot make statistical conclusions based on the transformed beta coefficients we will analyze in this study.

Another mode of analysis we will perform is a sub analysis of regression trees called bagging. Regression trees are an important median for predictive modeling in statistics where we partition a dataset into small groups, and fit regression models for each group. They allow us to highlight the most important variables in making predictions, and thus deliver another form of coefficient ranking in the job burnout and secondary trauma models. Bagging in the bootstrap aggregating technique of regression trees in which we combine and average many models for each subgroup. This technique reduces variability and overfitting, improving predictive performance.

Next, we will create subgroups of our observations using k-means clustering. Clustering allows us to identify observations that exhibit similar traits, and categorize these observations within each cluster. This analysis will help us determine whether or not our observations can be split by their responses to adaptive and maladaptive items.

Results

Below are the results of our multiple linear regression models, divided into the beta coefficients for job burnout and secondary trauma, ranked in descending order:

Beta Coefficients for Coping Strategies (Secondary Trauma)

Religious Coping	0.37003967
Emotional Support	0.35492847
Mental Disengagement	0.27631275
Substance	0.21561372
Positive	0.13743198
Behavioral Disengagement	0.13425438
Suppression	0.11963933
Active	0.08092615
Humor	0.01594904
Denial	-0.08211260
Acceptance	-0.11703396

Beta Coefficients for Coping Strategies (Job Burnout)

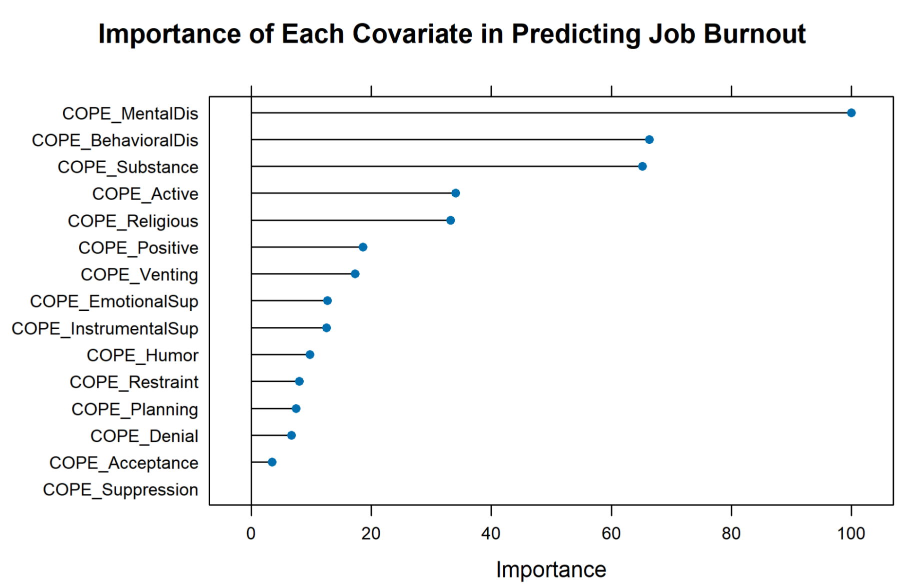
Mental Disengagement	0.442694359
Planning	0.321813079
Substance	0.245493802
Behavioral Disengagement	0.174944263
Emotional Support	0.120066341
Acceptance	0.088639393
Venting	0.064597243
Restraint	0.018160864
Humor	-0.003208883
Instrumental Support	-0.005968596

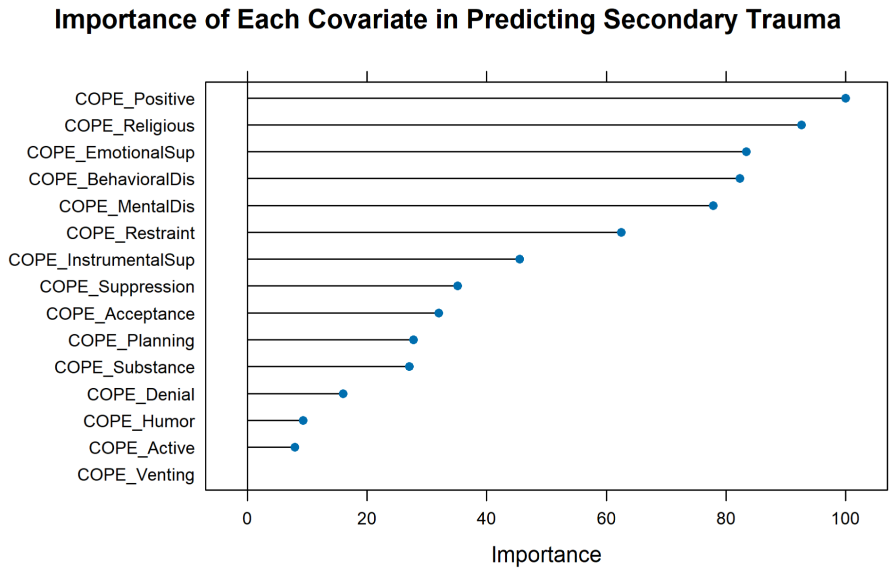
Planning	-0.12169143
Restraint	-0.16845572
Venting	-0.19544324
Instrumental support	-0.28318711

Religious	-0.074065286
Denial	-0.097582245
Positive	-0.184740904
Suppression	-0.192472118

Though we can’t interpret these coefficients from a scientific or psychological perspective, we can conclude that coping strategies with highly positive or highly negative scores, like mental disengagement in the job burnout model and instrumental support in the secondary trauma group, are more relevant to the regression model than coping strategies with coefficients closer to zero.

Next, we shift to our bagging regression trees to analyze the importance of each covariate in our two models. After creating bagging functions for each response variable, we create importance charts that highlight the variable importance across each bagged tree, which we visualize below:





These charts provide a scope into the importance of each coping strategy in predicting burnout and secondary trauma. While we cannot make any scientific conclusions on the role that each of these coping strategies plays, we can make distinct claims about specific covariates with respect to the model; For example, mental disengagement plays a more important role in our job burnout model than suppression.

Lastly, we move to k-means clusters to analyze the difference between our observations, with emphasis on adaptive and maladaptive traits. After creating a k-means cluster function, splitting the observations into two groups, we extract the mean scores of each coping strategy from the two clusters and calculate the difference between each variable. Below are the results of this calculation:

Positive: -0.1555556	**Mental Disengagement:** -0.5444444	**Venting:** -0.2777778	**Instrumental Support:** -1.3444444
Active: -0.3888889	**Denial:** -1.2444444	**Religious:** -6.8888889	**Humor:** 0.6666667
Behavioral Disengagement: -0.6000000	**Restraint:** -2.5666667	**Emotional Support:** -1.1666667	**Substance:** 0.2000000
Acceptance: 0.6888889	**Suppression:** -1.1666667	**Planning:** -1.1444444	

The majority of the coping strategies scores display minor differences between the two clusters, but religion stands out as a major difference. Considering religion is identified as an adaptive trait, we can conclude that the two clusters display differences in how the observations in the study utilized an adaptive coping strategy to combat compassion fatigue. With that said, we cannot make any major distinctions between how the observations utilized all adaptive or maladaptive coping strategies, as there is no evidence to suggest that the k-means analysis split our observations into groups with similarities across all adaptive or maladaptive strategies.

Discussion

To conclude, we were able to conduct linear regression, regression tree analysis, and k-means analysis to provide results for coping strategies and their relationship to compassion fatigue, as well as how the observations can be categorized by coping strategies.

Our first research objective tasked us with ranking our coping strategies with respect to compassion fatigue. The beta coefficients highlight how certain coping strategies play a more significant role with respect to compassion fatigue, but it doesn't provide a comprehensive rank, as highly positive and negative coefficients may have similar effects on the model. The bagging analysis, on the other hand, provides a more complete model in analyzing which covariates are most and least impactful in each respective model. It allows us to support what we already know on the subject. For example, Mental disengagement was the most important covariate in the job burnout model, which would support the idea that using mental disengagement as a coping strategy can be attributed to more extreme bouts of job burnout.

Ultimately, the limitations involved with this objective were our limited knowledge of beta coefficients, as well as our lack of scientific knowledge to connect our coping strategies and fatigue scores. Beta coefficients are outside the realm of statistical practice, so statistical conclusions would be inappropriate, and making any scientific claims on the covariates should be left to professionals who better understand these relationships.

Our second research objective had us focus on our observations, and whether or not they could be differentiated based on adaptive and maladaptive coping strategies. Our k-means analysis created two groups of observations, but provided no evidence to suggest that one group utilized adaptive or maladaptive strategies more potently. Ultimately, they could be differentiated by religious coping, leading us to conclude that strong differences in how religion is used as a coping strategy can be identified between the observations. The limitations in this k-means analysis pertain to the fact that we didn’t utilize demographic information to differentiate between our observations. If we included age, for example, we could get a better idea of how observed TCTSY-F’s differ between the two clusters. That said, using demographic info to predict responses in psychological studies is considered malpractice, and thus was not included in our study.

References

Nguyen-Feng, V. (n.d.). *Compassion Fatigue & Coping*. *Trauma Center Trauma Sensitive Yoga Study*.

Regression trees. Regression Trees · UC Business Analytics R Programming Guide. (n.d.). https://uc-r.github.io/regression_trees#bag

K-means cluster analysis. K-means Cluster Analysis · UC Business Analytics R Programming Guide. (n.d.). https://uc-r.github.io/kmeans_clustering

Appendix

Objective #1:

```{r}
data.1 <- na.omit(Data_STAT4893W[, 161:181])
data <- Data_STAT4893W[, 1:85]
```

```{r}
## Multiple Linear Regression

m1 <- lm(CFS_SecondaryTrauma_ln ~ COPE_Positive + COPE_MentalDis + COPE_Venting +
COPE_InstrumentalSup + COPE_Active + COPE_Denial + COPE_Religious + COPE_Humor +
COPE_BehavioralDis + COPE_Restraint + COPE_EmotionalSup + COPE_Substance +
COPE_Acceptance + COPE_Suppression + COPE_Planning, data = data.1)

beta.m1 <- lm.beta(m1)

m2 <- lm(CFS_JobBurnout_ln ~ COPE_Positive + COPE_MentalDis + COPE_Venting +
COPE_InstrumentalSup + COPE_Active + COPE_Denial + COPE_Religious + COPE_Humor +
COPE_BehavioralDis + COPE_Restraint + COPE_EmotionalSup + COPE_Substance +
COPE_Acceptance + COPE_Suppression + COPE_Planning, data = data.1)

beta.m2 <- lm.beta(m2)

var_names.1 <- names(coef(beta.m1))
coef_vals.1 <- coef(beta.m1)
beta.table.1 <- data.table::data.table("Coping Strategies" = var_names.1, "Betas" = coef_vals.1)
beta.table.1[order(coef_vals.1, decreasing = TRUE)]

var_names.2 <- names(coef(beta.m2))
coef_vals.2 <- coef(beta.m2)
beta.table.2 <- data.table::data.table("Coping Strategies" = var_names.2, "Betas" = coef_vals.2)
beta.table.2[order(coef_vals.2, decreasing = TRUE)]
```

```{r}
## Bagging with Caret
```

ctrl <- trainControl(method = "cv", number = 10)

CV bagged model

bagged_cv.1 <- train(CFS_SecondaryTrauma_ln ~ COPE_Positive + COPE_MentalDis + COPE_Venting + COPE_InstrumentalSup + COPE_Active + COPE_Denial + COPE_Religious + COPE_Humor + COPE_BehavioralDis + COPE_Restraint + COPE_EmotionalSup + COPE_Substance + COPE_Acceptance + COPE_Suppression + COPE_Planning, data = data.1, method = "treebag", trControl = ctrl, importance = TRUE)

plot(varImp(bagged_cv.1), 15, main = "Importance of Each Covariate in Predicting Secondary Trauma")

bagged_cv.2 <- train(CFS_JobBurnout_ln ~ COPE_Positive + COPE_MentalDis + COPE_Venting + COPE_InstrumentalSup + COPE_Active + COPE_Denial + COPE_Religious + COPE_Humor + COPE_BehavioralDis + COPE_Restraint + COPE_EmotionalSup + COPE_Substance + COPE_Acceptance + COPE_Suppression + COPE_Planning, data = data.1, method = "treebag", trControl = ctrl, importance = TRUE)

plot(varImp(bagged_cv.2), 15, main = "Importance of Each Covariate in Predicting Job Burnout")
```

Objective #2:

data.2 <- data.1[, 1:15]

```{r}
set.seed(1234)

k1 <- kmeans(data.2, centers = 2)
fviz_cluster(k1, data = data.2)
cluster.1 <- data.table::data.table("Variables" = names(data.2), "Cluster #1 Means" = k1$centers[1, ])
cluster.2 <- data.table::data.table("Variables" = names(data.2), "Cluster #2 Means" = k1$centers[2, ])

cluster.1[order(k1$centers[1, ], decreasing = TRUE)]
cluster.2[order(k1$centers[2, ], decreasing = TRUE)]

```
k1$centers[1, ] - k1$centers[2, ]
```