12-2017

Statistical linear mixed models for evaluation of training program in hand surgery chief residents

Zoe Michelle Ross
University of Tennessee at Chattanooga, jtd524@mocs.utc.edu

Statistical Linear Mixed Models for Evaluation of Training Program in Hand Surgery Chief Residents

Zoe Michelle Ross

Departmental Honors Thesis
The University of Tennessee at Chattanooga
Department of Mathematics

Examination Date: November 3rd, 2017

Dr. Cuilan Gao
Associate Professor of Statistcs
Thesis Director

Dr. Andrew Ledoan
Associate Professor of Mathematics
Department Examiner

Dr. Ossama Saleh
Professor of Mathematics
Department Examiner

Contents

Introduction

Background

Chief resident clinics (CRCs) are employed widely in the United States across various medical disciplines and thought to augment physician education while delivering quality medical care [1-2]. CRSs provide, "the opportunity for a surgeon to gain independent experience while offering cost effective benefits to patients," aim to allow residents to become primary care providers, build patient relationships, and follow through with plans of care [3]. Hand surgery CRCs are thereby thought to catalyze the achievement of milestones in patient care by providing greater liberty for operative autonomy and continuity of care. This educational value has received much attention in the literature for cosmetic CRCs [4-5]. However, reports are lacking for Hand CRCs, which may be even more relevant to graduating competent hand surgeons in common on-call scenarios. The program that is the focus of this study incorporates a half-day hand CRC into the weekly training curriculum that predominantly features general hand surgery patients with a focus on trauma and other emergency referral care. These Hand CRCs allow chief residents to make clinical decisions under the supervision of experienced surgery faculty members. The general philosophy is that residents should approach CRC patients as their primary providers with assistance from attending physicians provided to a degree that is proportional with the trainee's demonstrated level of competency and independence. Attending physicians ultimately take responsibility for all care delivered but provide significant autonomy to residents both in the clinical consultation and during operative intervention.

The definition of quality medical education has been a subject of investigation since the Flexner Report was introduced at start of the 20th century. The American Council of Graduate Medical Education (ACGME) continues to develop this concept despite criticism from surgical education programs for enacting the Institute of Medicine's duty hour restrictions [6-10].

Critics argue that achieving traditionally time-based competencies in surgical autonomy and continuity of patient care are threatened by duty hour restrictions. This has caused a paradigm shift toward "milestone" competency-based training, which requires documentation of progressive educational achievement.

Our program incorporates a half-day hand surgery chief resident clinic (H-CRC) into the weekly training curriculum, which cares predominantly for acute trauma or other emergency room referral patients. Chief residents in hand surgery thereby act as primary providers and make clinical decisions under the supervision of board-certified University of Tennessee College of Medicine (UTCOM) faculty members with Certificates of Added Qualification (CAQ) in hand surgery. Attending physicians ultimately take responsibility for all care delivered but allow significant autonomy in both the clinic consultation and operative intervention commensurate to the trainee's demonstrated level of competency and independence.

We demonstrate the achievement of ACGME patient care milestones in surgical autonomy within an H-CRC model. This is the first study to provide verification of competency-based hand surgery milestone achievement in an H-CRC. It establishes an evidence-based method for the documentation of residents' progress in hand surgical patient care, as outlined by the Hand Surgery Milestone Project and Orthopedic Surgery Milestone Project.

Following approval by the University of Tennessee Chattanooga Institutional Review Board, a retrospective review of all patients at Erlanger Hospital in H-CRC from October 1st, 2010 to October 1st, 2015 was collected. All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2008. 46 Approval to defer patient informed consent was granted by University of Tennessee Chattanooga institutional review board, given that data collection was retrospective and posed no risk to included patients.

Motivation

Resident clinics (RCs) are intended to catalyze the achievement of educational milestones through progressively autonomous patient care. However, few studies quantify their effect on competency-based surgical education, and no previous publications focus on hand surgery RCs.

Study Objectives

The objectives of my thesis are as follows:

1. To determine whether there is a statistically significant association between resident's performance on surgeries and the traning years.
2. To determine additional factors such as sugery type that affect the resident's performance on surgeries.
3. To determine if there is a random factor, such as residents that will affect the autonomy score.

Hypothesis

We hypothesize that the resident's performance (measured by autonomy scores) is positively associated with training years, and the resident is random factor of resident's performance.

Goals of the Study

This study aims to use statistical theories and knowledge of descriptive statistics and inference statistics, such as confidence intervals, two sample t-tests, correlation and association tests, as well as statistical model building such as analysis of variance with random effects and mixed linear models. We hypothesize that the higher the training years, the higher the autonomy score will be.

Description of the Data

Data Collection

Following approval by The University of Tennessee Chattanooga Institutional Review Board, A total of 826 patients at Erlanger Hospital from October 2010 to October 2015 was collected. All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national). Investigators compiled data on patient demographics, provider encounters, operational statitics, operative details, and dicated surgical autonomy score on an ascending 5 point scoring system.

Description of Variables

This data set contains 826 hand surgery patients, and each row of data represents one surgery case.

1. The main variables contained in this data set include patient information, resident information, and measures of performance of chief residents on surgeries. Below is a list of the variables: **Last Name**, **First Name**, **Date of Birth**, **Age**, **Race**, **Gender**, **Insurance Status**, **Clinic Appointment Date**, **Resident**, **Diagnosis**, **Procedure**, **Autonomy Score** , **Resident Training Level**, **Resident**, **Total Number of Appointments**, **Pre or Post OP Visit Same**, **Direct Cost**, **Payment**, **Indirect Cost**.

   Although there are many variables, this study mainly focuses on those variables that we believe are most associated with the **Autonomy Score** of the surgery. These variables are described below

2. The dependent variable of this study is the **Autonomy Score** which is a measure of how well a surgery was performed. A high **Autonomy Score** indicates that a surgery went well, and a low **Autonomy Score** indicates that a surgery did not go well.

3. The main independent variable of statistical model is the **Training Level**. A resident's **Training Level** indicates how many years they have been a resident training program. One indicates most experienced and 9 indicates least experienced.

4. Additional independent variable is the **Surgery Type**. There are ten different surgery types.

5. The next independent variable of this statistical model is the **Residents**. The residents each have different experience and training levels therefore different residents may be in consideration if the study is repeated.

Methodology

Summary of Statistical Methods

Descriptive statistics including means, standard deviations, and graphs were used to report the distribution of the autonomy score. Confident intervals are used to estimate the true means of outcomes of interest in each group. Simple linear regression models, a multiple linear regression model, and linear mixed models were developed to investigate how training training level affects the autonomy score for each resident. We also explored the other covariates such as surgery types and random effects.

Simple Linear Regression

The simple linear regression model is a model that allows us to provide a relationship between two variables– an independent variable, x, and a dependent variable, y, where x is a fixed effect factor. This model seeks to express the relationship between these two variable by deriving a best fitting line that may takes on the form,

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

where β_0 and β_1 represent the unknown regression parameters we seek to estimate and ϵ_i represents the error term, for $i = 1, 2, 3..., n$, the error terms are assumed to have constant variances, σ^2, such that the error terms are independent and identically distributed with $\epsilon_i \sim N(0, \sigma^2)$.The coefficients, β_0 and β_1, of The simple linear regression model can be obtained by using the ordinary least squares method (OLS), which seeks to minimize the the squared differences between the observed data points and the estimated to get a best approximation for the independent variable, $\mathbf{y}$.

Derivation of β_0 and β_1 Using Ordinary Least Squares Method

$Min \sum \epsilon_i^2 = Min \sum (y_i - \hat{y})^2$

Let the least squares function be $S = Min \sum (y_i - \beta_0 - \beta_1 x)^2$

To optimize, we take $\frac{\partial S}{\partial \beta_0}$ and $\frac{\partial S}{\partial \beta_1}$ separately and set each equal to zero,

$\frac{\partial S}{\partial \beta_0} = -2 \sum (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$

$\frac{\partial S}{\partial \beta_1} = -2 \sum (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) x_i = 0$

Simplifying we obtain,

$n\hat{\beta}_0 + \hat{\beta}_1 \sum x_i = \sum y_i$

$\hat{\beta}_0 \sum x_i + \hat{\beta}_1 \sum x_i = \sum y_i x_i$

Solving for β_0 and β_1,

$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

$\hat{\beta}_1 = \frac{\sum y_i x_i - \sum y_i (\sum x_i)/n}{\sum x_i^2 - (\sum x_i)^2/n} = \frac{S_{xy}}{S_{xx}}$

where $S_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y})$ and $S_{xx} = \sum (x_i - \bar{x})^2$

Multiple Linear Regression Model

The multiple linear regression model is an extension of the simple linear regression model. Its objectives are the same. However, the main distinction here is that it considers more than one independent, fixed effect variable. In algebraic form it can be expressed as

$$y_i = \beta_0 + \beta_1 x_i 1 + \beta_2 x_i 2 + ... + \beta_k x_i k + \epsilon_i$$

Due to this model having the nature of multiple independent variables, it is best to use the matrix form of the linear regression model,

$$Y = X\beta + \epsilon$$

where **X**, is called the **design matrix**, assumed to have full rank and the conditions around ϵ hold.

Derivation of the Ordinary Least Squares Coefficients in Matrix Form

We begin with

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon$$

Solving for ϵ we get,

$\epsilon = \mathbf{Y} - \mathbf{X}\beta$

Let $S = [\sum \epsilon^2] = \epsilon'\epsilon$ Then

$S = (\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)$

$= \mathbf{Y'Y} - \mathbf{YX'}\beta' - \mathbf{YX}\beta + \mathbf{X'}\beta\mathbf{X}\beta$

$= \mathbf{Y'Y} - \mathbf{2YX'}\beta' + \mathbf{X'}\beta\mathbf{X}\beta$

Taking the partial derivative of S with respect to β to optimize we obtain,

$\frac{\partial S}{\partial \beta} = -2\mathbf{X'y} + \mathbf{2}\hat{\beta}\mathbf{X'X} = \mathbf{0}$

Now solving for $\hat{\beta}$,

$-2\mathbf{X'Y} = -2\hat{\beta}\mathbf{X'X}$

Dividing both sides by -2 we obtain,

$\mathbf{X'Y} = \hat{\beta}\mathbf{X'X}$

Taking the inverse, $\mathbf{X'Y}(\mathbf{X'X})^{-\mathbf{1}} = \hat{\beta}(\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'X}$

Finally, we see that we get the $\hat{\beta}$ with minimum error,

$\hat{\beta} = (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'Y}$.

Mixed Linear Regression Model

The linear mixed regression model (LMM) is also an extension of a simple linear regression model, but with an extra complexity called a random effect. The LMM may consider one or more fixed effect factors as well as one or more random effect factors. The random factor usually is a qualitative variable whose levels are random samples from a population or level being studied. For example, consider an experiment to investigate the effect of several drug treatments on sample of patients. Typically, we are interested in specific drug treatments, so we would treat the drug effect as fixed. However, it makes most sense to treat the patients effects as random. Because it reasonable to treat the patients as being randomly selected from a larger collection of patients whose characteristics we would like to estimate. Furthermore we are not particularly interested in these specific patients, but in the whole population of patients. A random effects approach to modeling effects is more ambitious in the sense that it attempts to say something about the wider population beyond the particular sample. To account for this, the LMM considers the error term ϵ differently than it would in a simple model. The random effects will give structure to ϵ. The model can account for individual differences by assuming different random intercepts for each subject, which the model will estimate. Similar to the simple linear regression model, the LMM can be expressed in algebraic form or matrix form.

Algebraic Form

$Y_i = \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_p X_{ip}$ (fixed)$+ u_{1i} Z_{i1} + ... + u_{qi} Z_{iq} + \epsilon_i$ (random)

where the index, i, is used to index subjects. Note that $Z_{i1}, ..., Z_{iq}$ is associated with the random effects $u_{1i}, ..., u_{qi}$ that are specific to subject i.

Since the LMM can take on one or more fixed or random effects, it is best to use the matrix form of this equation.

Matrix Form

$\mathbf{Y_i} = \mathbf{X_i}\beta$ (fixed)$+\mathbf{Z_i}\mathbf{u_i} + \epsilon_i$ (random)

where $u_i \sim N(\mathbf{0}, \mathbf{D})$ and $\epsilon \sim N(\mathbf{0}, R_i)$

and $\mathbf{D}$ represents variance of $\mathbf{u_i}$ and $\mathbf{R_i}$ represents variance of $\epsilon_\mathbf{i}$

Results

Autonomy Scores Across Different Levels of Training Years

The boxplot shows how the **autonomy scores** vary by groups of **training levels**. Generally speaking, a resident with longer training period has a higher autonomy scores. For example, the residents with eight years training has higher average **autonomy score** than the residents with one year training.

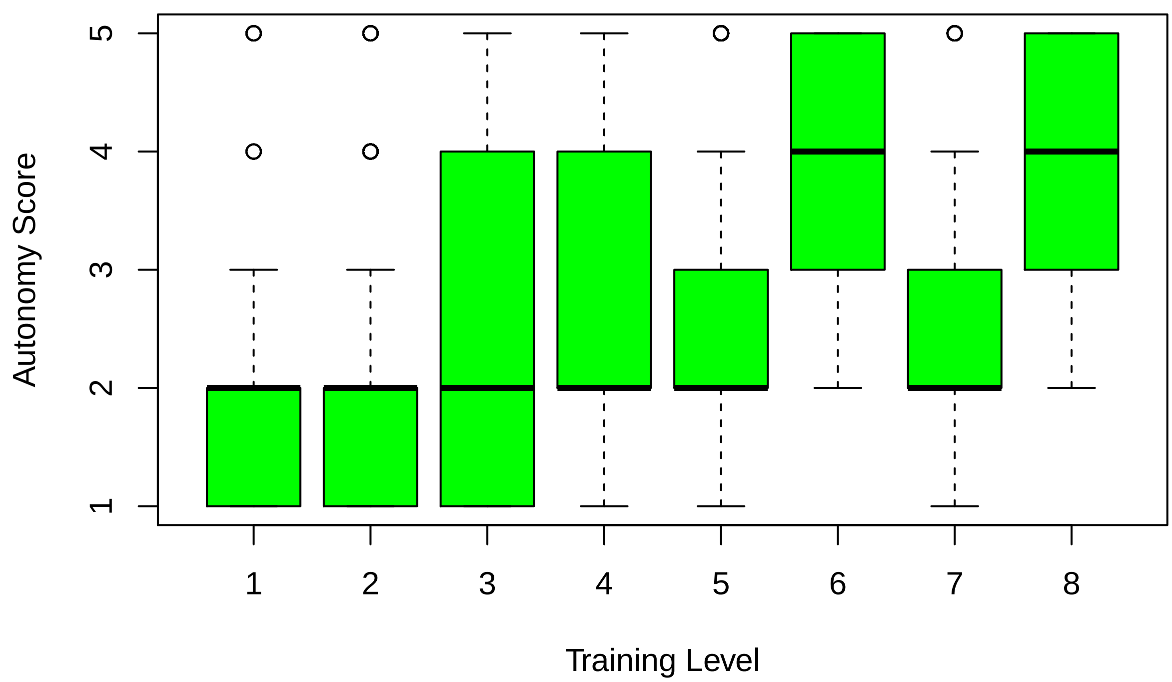

Figure 1: Distribution of Autonomy Scores for Each Training Level

Pairwise Comparison of Autonomy Scores between Different Training Levels

A pairwise comparison is a an Analysis of Variance (ANOVA) technique used in multiple comparison scenarios. This technique allows us to look at all pairs of levels for a factor. In our case, we can observe all pairs of the training years. From the table, we can see the confidence interval and the p-value for each pair of training years. This allows us to view any significant differences in autonomy score between the training years. We may note that the first training years with a significant difference is between residents with one year of training and residents with three years of training. This means that residents with three years of training have significantly higher autonomy scores than residents with one year of training. Furthermore, we can see that the most significant differences are between residents with eight years of training and residents with one or two years of training. These results can also be seen in the 95% family-wise confidence level graph. The horizontal lines represents the confidence interval for each pair of training levels. The vertical line at $x = 0$ represents zero difference between the two groups. Therefore, we have strong evidence to support that the confidence intervals for the training level pairs that do not include zero, have significantly different autonomy scores.

Table 1: Pairwise Comparison of Autonomy Scores Among Training Levels

	diff	lwr	upr	p adj
2-1	0.0351917	-0.5141199	0.5845033	0.9999995
3-1	0.5774006	0.0347394	1.1200618	0.0277672
4-1	0.8930348	-0.3298122	2.1158818	0.3409458
5-1	0.6960651	0.1185135	1.2736168	0.0064622
6-1	1.9347015	0.4754342	3.3939688	0.0015794
7-1	0.6799547	0.0320485	1.3278608	0.0318707
8-1	1.9687924	0.6996798	3.2379050	0.0000782
3-2	0.5422089	0.1654604	0.9189574	0.0003695
4-2	0.8578431	-0.3009550	2.0166412	0.3230782
5-2	0.6608734	0.2354041	1.0863428	0.0000760
6-2	1.8995098	0.4934802	3.3055394	0.0011561
7-2	0.6447630	0.1278098	1.1617161	0.0040204
8-2	1.9336007	0.7260805	3.1411209	0.0000378
4-3	0.3156342	-0.8400262	1.4712946	0.9913766
5-3	0.1186645	-0.2981833	0.5355124	0.9889464
6-3	1.3573009	-0.0461439	2.7607456	0.0663832
7-3	0.1025540	-0.4073269	0.6124350	0.9987330
8-3	1.3913918	0.1868823	2.5959013	0.0111030
5-4	-0.1969697	-1.3694183	0.9754789	0.9996102
6-4	1.0416667	-0.7389318	2.8222651	0.6351589
7-4	-0.2130802	-1.4217373	0.9955770	0.9994641
8-4	1.0757576	-0.5526517	2.7041668	0.4770682
6-5	1.2386364	-0.1786646	2.6559373	0.1379169
7-5	-0.0161105	-0.5629772	0.5307563	1.0000000
8-5	1.2727273	0.0521013	2.4933532	0.0339569
7-6	-1.2547468	-2.7021439	0.1926502	0.1451024
8-6	0.0340909	-1.7785930	1.8467748	1.0000000
8-7	1.2888377	0.0333919	2.5442836	0.0393855

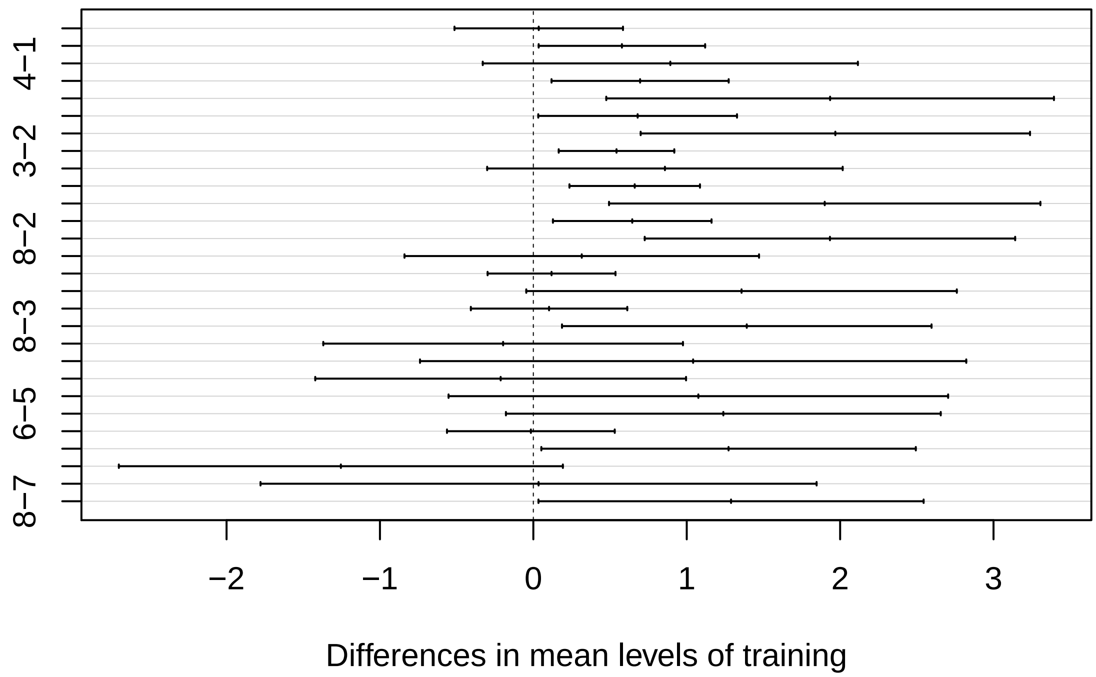

Figure 2: 95% Family-wise Confidence Level

Result of Simple Linear Regression Model

The first model that was built was a simple linear regression model that was used to measure the autonomy score based off of one fixed factor, training level. This regression takes the form of $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, which was derived above. We define β_0 to be the intercept, or the base autonomy score if the resident had zero years of training. β_1 is the regression coefficient that represents the mean change in autonomy score for every one unit increase of x_i, or training level. ϵ_i is the error, or residuals.y_i is the dependent variable, autonomy score.

The equation we obtain is $\hat{y} = 3.30 - 0.16x$ From the model we can observe that for every one unit "increase" in training level, the autonomy score will fall by 0.16, so this is the coefficient for the resident training level. Note that this negative relationship is due to training level 1 being most experienced.

Table 2: Result of Simple Linear Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.3028174	0.15140476	21.814488	0e+00
Training Level	-0.1648041	0.02601318	-6.335408	4e-10

Result of Multiple Linear Regression Model

The next model that was built was a multiple linear regression model that was used to measure the autonomy score based off of two fixed factors, training level and surgery type. This regression takes the form of $y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i + ... + \beta_n x_i + \epsilon_i$. We define β_0 to be the intercept, or the base autonomy score if the resident had zero years of training. β_1 is the regression coefficient that represents the mean change in autonomy score for every one unit increase of x_i, or training level. β_2 is the regression coefficient that represents the mean change in autonomy score for every one unit increase of x_i, or surgery type. ϵ_i is the error, or residuals y_i is the dependent variable, autonomy score.

The equation we obtain is $\hat{y} = 0.291 - 0.15x_1 + 0.06x_2$ From the model we can observe that for every one unit "increase" in training level, the autonomy score will fall by 0.16, so this is the coefficient for the resident training level. Note that this negative relationship is due to training level 1 being most experienced. We can note that each variable is significantly associated with the autonomy score. The P-values of the training levels and surgery type are $P = 0.1807$ and $P = 0.000026$, respectively. This indicates that each variable plays a significant role in measuring the autonomy score.

Table 3: Result of Multiple Linear Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	2.91315333	0.22194602	13.12550	0.0000000000
Surgery Type	0.05775301	0.02432658	2.37407	0.0180668667
Training Level	-0.14637744	0.03440411	-4.25465	0.0000261318

Result of Mixed Linear Regression Model

Mixed Linear Regression Model for One Fixed Factors

To enhance the model further, we consider a mixed linear regression model where the resident who performed the surgery is the radom factor. We have the equation $Y_{ti} = \beta_1 \times X_{ti}^1 + (fixed) + u_{1i} \times Z_{ti}^1 + \epsilon_{ii}(random)$ where β_1 is represents the mean change in autonomy score for every one unit increase in training level and $u_{1i} \times Z_{ti}^1$ represents the random factor, residents.

Now, consider the results of the first mixed linear regression model. The fixed effect can be interpreted as before. We see that training years is significantly associated with the autonomy score, with the P-value= 0.000000489 indicating that training level significantly impacts the autonomy score. Now, looking at the random effect results, we must note the standard of deviation column, which equals 0.2902 for residents. This tells us how much variability in the autonomy score is due to different residents performing the surgery. Lastly, we have the residuals. This is how much of the autonomy score is not explained by our model. Here we have that the residuals equals 1.274.

Table 4: Result of Mixed Linear Model with one Fixed Factor (Training Level) and random factor (Residents)

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	3.3711531	0.1836641	116.6526	18.354992	0e+00
Traning Level	-0.1747904	0.0326398	107.5957	-5.355129	5e-07

Mixed Linear Regression Model for Two Fixed Factors

Next, we take the considered the second mixed linear regression model with two fixed factors, training level and surgery type, and the same random factor as the previous model. Consider the results of the second mixed linear regression model. The fixed effects can be interpreted as before. With this model, we see that both fixed effects, training level and surgery type are significantly related to the autonomy score, with P=0.0204 and P=0.0003, respectively. We see that the random factor, residents has a standard of deviation equal to 0.3689. Here we have that the residuals equals 1.1894.

Table 5: Result of Mixed Linear Model with Two Fixed Factor (Training Level and Surgery Type) and one random factor (Residents)

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	3.0175864	0.2577901	149.0980	11.705593	0.0000000
Surgery Type	0.0559634	0.0240270	396.3339	2.329186	0.0203510
Traning Level	-0.1606913	0.0425839	109.6414	-3.773527	0.0002615

Conclusion

By analyzing a simple linear regression model, we found evidence to support that the autonomy score is significantly impacted by the training years of a resident. Secondly we considered a multiple linear regression model that determined how both training years and surgery type affect the autonomy score. Lastly, we considered two linear mixed regression models, with residents being the random factor and with and without the surgery type in the model. By comparing all the models, we found that the mixed linear regression model with fixed factors: training level and surgery type and one random factor: residents, appears to be the best model to predict the autonomy score. In this model, all of the variables are significantly associated with the autonomy score. Furthermore, this model is associated with the lowest residuals. Therefore, we conclude that the progressive autonomy score of a resident on a hand surgery depends on both the training levels of the resident and the type of surgery that the residents conducted. In addition, the performance may also vary from resident to resident given the same training level and surgery type.

Our model enables the hand surgery trainees to practice progressively autonomous hand surgery care and provide a formula for training programs to observe progressive autonomy. Hand residents clinics thereby provide a mechanism for individual and program self-assessment in the demonstration of progressive operative autonomy in hand surgery. We believe that hand resident clinics can be valuable for increasing independent surgical decision making and operative proficiency, though further studies are needed to further define the quality of care delivered and their role in other competency-based curriculums.

Acknowledgements

I would first like to thank Dr. Gao, as she has provided me with endless mentorship, encouragement, and patience. Completion of the thesis would not have been possible without her prompt and constructive feedback. I also want to thank Dr. Ledoan for suggesting that students pursue an undergraduate thesis and for his willingness to serve on my thesis committee. I am also thankful for the additional support from Dr. Saleh. Although I have never been his student, he generously agreed to serve on the thesis committee.

References

1. Fox J., Weisberg, S. (2011) Cox proportional-hazards regression for survival data in R. An appendix to an R companion to applied regression, 2nd edition.
2. Cooke M., Irby D.M., Sullivan W., Ludmerer K.M., American medical education 100 years after the Flexnar report. N Engl J Med. 2006;355:1339-1344.
3. Flexnar A., Medical Education in the United States and Canada: A Report to the Carnegie Foundation for the Advancement of Teaching. New York: Carnegie Foundation for the Advancement of Teaching: 1910. 3. Graduate Medical Education. Report of the Commission on Graduate Medical Education. Chicago: University of Chicago Press. 1940.
4. Ahmed N., Devitt K.S., Keshet I. et al. A systematic review of the effects of resident duty hour restrictions in surgery. Ann Surg. 2014;259:1041-1053.
5. Basu C.B., Chen L.M., Hollier L.H. Jr., Shenaq S.M., The Effect of the Accreditation Council for Graduate Medical Education Duty Hours Policy on plastic surgery resident education and patient care: an outcomes study. Plast Reconstr Surg. 2004 Dec;114(7):1878-86.
6. Bilimoria KY, Chung J.W., Hedges L.V., Dahlke A.R., Love R., Cohen M.E., Tarpley J., Mellinger J., Mahvi D.M., Kelz R.R., Ko C.Y., Hoyt D.B., Lewis F.H. Development of the Flexibility in Duty Hour Requirements for Surgical Trainees (FIRST) Trial protocol: a national cluster-randomized trial of resident duty hour policies. JAMA Surg. 2016 Mar 1;151(3):273-81.
7. Durkin E.T., McDonald R., Munoz A., Mahvi D., The impact of work hour restrictions on surgical resident education. J Surg Educ. 2008;65:54-60.
8. Institute of Medicine, Resident Duty Hours: Enhancing Sleep, Supervision, and Safety. Washington, D.C.: The National Academies Press; 2009.
9. Nasca T.J., Day S.H., Amis E.S. Jr., ACGME Duty Hour Task Force. The new recommendations on duty hours from the ACGME Task Force. N Engl J Med. 2010;363:e3.
10. Philibert I., Friedmann P., Williams W.T., ACGME Work Group on Resident Duty Hours. Accreditation Council for Graduate Medical Education. New requirements for resident duty hours. JAMA. 2002;288:1112-1114.

Appendix of R codes

```
# Hdata=read.table("C:/Users/Zoe-Honor Thesis/hand.data.txt", sep=",", header=F, strin
# colnames(Hdata)=NULL
# colnames(Hdata)=Hdata[1, ]
# Hdata=Hdata[-1, ]
# attach(Hdata)
# Hdata$AUTONOMY.LEVEL=as.numeric(Hdata$AUTONOMY.LEVEL)

# mean autonomy score (how well the surgery was performed)
# mean(Hdata$AUTONOMY.LEVEL, na.rm=T)
# colnames(Hdata)[1:44]= c("MRN", "Lname", "Fname","DOB", "age", "race", "gender", "in

# dim(Hdata) # 825 cases
# colnames(Hdata)
# auto.hand=as.numeric(Hdata$AUTONOMY.LEVEL)
# summary(auto.hand)  # 72 NAs
# number of rows
# m=nrow(Hdata)
# print(m)
# resid.hand=NULL
# fixing the names of the residents
# for(i in 1:m)
# {resid.hand[i]=unlist(strsplit(Hdata$RESIDENT.1[i], split=","))[1]}

# unique(resid.hand)
# resid.hand[resid.hand=="suthpin"]<-"SUTPHIN"
# resid.hand[resid.hand=="Sutphin"]<-"SUTPHIN"
# resid.hand[resid.hand=="Davit"]<-"DAVIT"
# resid.hand[resid.hand=="davit"]<-"DAVIT"
# resid.hand[resid.hand=="Flavia Davit"]<-"DAVIT"
# resid.hand[resid.hand=="Fernandez"]<-"FERNANDEZ"
# resid.hand[resid.hand=="Mooty"]<-"MOOTY"
# resid.hand[resid.hand=="higgins"]<-"HIGGINS"
# resid.hand[resid.hand=="spitler"]<-"SPITLER"
# resid.hand[resid.hand=="Spitler"]<-"SPITLER"
# resid.hand[resid.hand=="Jesse Doty"]<-"DOTY"
# resid.hand[resid.hand=="doty"]<-"DOTY"
# resid.hand[resid.hand=="Doty"]<-"DOTY"
# resid.hand[resid.hand=="Griner"]<-"GRINER"
# resid.hand[resid.hand=="griner"]<-"GRINER"
# resid.hand[resid.hand=="Phillips"]<-"PHILLIPS"
# resid.hand[resid.hand=="jarrell"]<-"JARRELL"
# resid.hand[resid.hand=="Jarrell"]<-"JARRELL"
```

```
# resid.hand[resid.hand=="Landis"]<-"LANDIS"
# resid.hand[resid.hand=="Bruce"]<-"BRUCE"
# resid.hand[resid.hand=="Sweets"]<-"SWEETS"
# resid.hand[resid.hand=="Howell"]<-"HOWELL"
# resid.hand[resid.hand=="Dunn"]<-"DUNN"
# resid.hand[resid.hand=="Lemons"]<-"LEMONS"
# resid.hand[resid.hand=="Pankiw"]<-"PANKIW"
# resid.hand[resid.hand=="Jackson"]<-"JACKSON"  # is the same as chapman-jackson
# resid.hand[resid.hand=="Fogleman"]<-"FOGLEMAN"
# resid.hand[resid.hand=="Nair"]<-"NAIR"
# resid.hand[resid.hand=="Rau"]<-"RAU"
# resid.hand[resid.hand=="STEFFEN"]<-"STEFFAN"
# resid.hand[resid.hand=="chapman-jackson"]<-"CHAPMAN" # correct?
# resid.hand[resid.hand=="Boaen"]<-"BOAEN"
# resid.hand[resid.hand=="cowart"]<-"COWART"
# resid.hand[resid.hand=="Cowart"]<-"COWART"
# resid.hand[resid.hand=="Dale"]<-"DALE"
# resid.hand[resid.hand=="#N/A"]<-"NA"
# resid.hand[resid.hand=="#REF!"]<-"NA"
# resid.hand[resid.hand=="???"]<-"NA"
# sort(unique(resid.hand))
# table(resid.hand)
#
# hand.data=Hdata
# hand.data$resid.hand=resid.hand
# attach(hand.data)

# Use hand.data for the fowllowing r code
# look at the distribution of autonomy level
# hand.data$AUTONOMY.LEVEL=as.numeric(hand.data$AUTONOMY.LEVEL)
# plot(density(hand.data$AUTONOMY.LEVEL, na.rm = T))

# look at dist. of autonomy modified
# Hdata$AUTONOMY.LEVEL=as.numeric(Hdata$AUTONOMY.LEVEL)
# plot(density(Hdata$AUTONOMY.LEVEL, na.rm = T))
# The data is skewed and not normally distributed.

# calculate the score of autonomy level for different resident
# AUTONOMY.LEVEL=as.numeric(AUTONOMY.LEVEL)
# mean(AUTONOMY.LEVEL, na.rm=T)
# Hdata[Hdata$resid.hand==SUTPHIN]
# Hdata[Hdata$resid.hand== MOOTY]

# This line of code successfully produces the mean AUTONOMY.LEVEL for each resident
```

```
# means=tapply(Hdata$AUTONOMY.LEVEL, Hdata$resid.hand, mean, na.rm=T)

# This line of code successfully computes the standard of deviation of the autonomy le
# std=tapply(Hdata$AUTONOMY.LEVEL, Hdata$resid.hand, sd, na.rm=T)

# The following lines of code make a table with columns of mean and std, rows of resid
# cbind(mean=tapply(Hdata$AUTONOMY.LEVEL, Hdata$resid.hand, mean, na.rm=T),std=tapply(
# At this point, we have a table with the mean and std of the autonomy score for each


# Part2 now calculate the sample mean, sample sd and CI ( 95% ) for groups by differen

# This code successfully calculates the sample mean and standard deviation for groups
# res1=cbind.data.frame(mean=tapply(Hdata$AUTONOMY.LEVEL, Hdata$RESIDENT.PGY.LEVEL..OP

# row.names(res1)=NULL
# print(res1)
#From above we have a table with the mean and standard of deviation for the autonomy s

# Now calculating the 95% CI.
# This efficiently calculates the LB and UB for the .95 confidence interval for groups
# sd=tapply(Hdata$AUTONOMY.LEVEL, Hdata$RESIDENT.PGY.LEVEL..OP, sd, na.rm=T)
# mean=tapply(Hdata$AUTONOMY.LEVEL, Hdata$RESIDENT.PGY.LEVEL..OP, mean, na.rm=T)
# ss= table(Hdata$RESIDENT.PGY.LEVEL..OP)
# note that ss=sample size
# zc=qnorm(.975)
# note that the zc is the quantile. We use .975 because we have 95% CI and this is two
# err=zc*(sd/ss^.5)
# LB=mean-err
# UB=mean+err

#Table with level corrected for 8 being most experienced below
# res1=cbind(level=c(8, 7, 6, 5,4,3,2,1, "u"), res1, cbind(LB, UB))
# library(knitr)
# #print(res1)
# kable(res1)
#Now we have a table with training level, mean, std, LB, and UB for residents by level


#Part 3: compare the significant difference between level 8 and level 1 using t-test

# Hdata$AUTONOMY.LEVEL[Hdata$RESIDENT.PGY.LEVEL..OP == 8]
# Hdata$AUTONOMY.LEVEL[Hdata$RESIDENT.PGY.LEVEL..OP == 1]
# t.test(Hdata$AUTONOMY.LEVEL[Hdata$RESIDENT.PGY.LEVEL..OP == 8],Hdata$AUTONOMY.LEVEL[
```

```
# t.test(Hdata$AUTONOMY.LEVEL[Hdata$RESIDENT.PGY.LEVEL..OP == 8],Hdata$AUTONOMY.LEVEL[

#From the t-test above, we can see that there is a significant difference as for both
#From the t-test above, we can see that there is a significant difference between 1 ye

# Part 4: Multiple groups comparison ( look at the overall difference among all the 8
#multiple groups comparison using ANOVA

# Hdata$RESIDENT.PGY.LEVEL..OP=as.factor(Hdata$RESIDENT.PGY.LEVEL..OP)
# res.aov=aov(AUTONOMY.LEVEL~RESIDENT.PGY.LEVEL..OP, data=Hdata)
# summary(res.aov)
# Simple Linear Regression with one factor (training level) (we can treat this model a
# Hdata$RESIDENT.PGY.LEVEL..OP=as.numeric(Hdata$RESIDENT.PGY.LEVEL..OP)
# lm1=lm(AUTONOMY.LEVEL~RESIDENT.PGY.LEVEL..OP, data=Hdata)
# sum.lm1=summary(lm1)
# print(sum.lm1)
# four decimals for p-value , and two decimals for rest
# kable(sum.lm1$coefficients)
# digits = c(2, 2, 2, 4), caption = "Table of Estimates of SLM")

#We can observe that for every one "increase" in training level, the autonomy score wi

# how the surgery type affects the autonomy score
# cleaning up the data for the surgery type. There are 10 types.
# t1=c(15100, 26605, 26607, 26608, 26705, 26615)
# t2=c(10180, 12044, 15002, 20103, 10060, 10061, 10180, 25028, 26011, 26034, 26025,110
# t3=c(15240, 25270, 25280, 25310, 26410, 26418, 26420, 26428, 26432, 26433)
# t4=c(25260, 25270, 25272, 25280, 25310, 25505, 26350, 26352, 26370, 26440)
# t5=c(26356, 26357)
# t6=c(10120, 10121, 26075, 20520, 20525, 20670, 20680, 20694, 24200, 25248, 26320)
# t7=c(25111, 25112, 25071, 26160, 26200)
# t8=c(25431, 25645, 25440, 25628, 25645, 25685)
# t9=c(26055)
# t10=c(25607, 25350, 25360, 25545, 25606, 25608, 25609)

# proc2=list(t1=t1, t2=t2, t3=t3, t4=t4, t5=t5, t6=t6, t7=t7, t8=t8, t9=t9, t10=t10)
# type2=rep(NA, nrow(Hdata))

# for(i in 1:10)
# { for( j in 1:nrow(hand.data))
#   {
#   if (is.element(Hdata$CPT[j], proc2[[i]]))
#     type2[j]=names(proc2[i])
#   }
```

```
# }

# table(type2)
# Hdata$Type=type2

#The table below will tell how many observations there are for each surgery type
# table(Hdata$Type)
# Hdata$Type=as.numeric(Hdata$Type)

# res.type=lm(AUTONOMY.LEVEL~Type, data=Hdata)
# summary(res.type)
# print(summary(res.type))
# Note that when interpreting this linear regression, type one is excluded from the ou
# multiple regression model with two fixed effects ( training level + surgery type(fac

# Ensuring the variables are read as numeric (Not strings)
# Hdata$Type=as.numeric(Hdata$Type)
# Hdata$RESIDENT.PGY.LEVEL..OP=as.numeric(Hdata$RESIDENT.PGY.LEVEL..OP)
# successful multiple regressions with both training level AND type
# res.pgytype=lm(AUTONOMY.LEVEL~Type+RESIDENT.PGY.LEVEL..OP, data=Hdata)
# summary(res.pgytype)

#Note that the above model uses type 1 (t1) as a base. So the values that we see from
# multiple regression model with two fixed effects (training level + surgery type(code
# Type=Hdata$Type
# Type[Type=="t1"]=1
# Type[Type=="t2"]=2
# Type[Type=="t3"]=3
# Type[Type=="t4"]=4
# Type[Type=="t5"]=5
# Type[Type=="t6"]=6
# Type[Type=="t7"]=7
# Type[Type=="t8"]=8
# Type[Type=="t9"]=9
# Type[Type=="t10"]=10
# Hdata$coded.Type=as.numeric(Type)

# lm2=lm(AUTONOMY.LEVEL~coded.Type+RESIDENT.PGY.LEVEL..OP, data=Hdata)
# sum.lm2=summary(lm2)
# print(sum.lm2)
# kable(sum.lm2$coefficients)

# THe mixed model with two fixed factors ( traning level and type of surgery ) and ran
# Download the package
```

```r
# library(lmerTest)
# LMM regression with fixed factors of autonomy level and training type and the fixed
# LMM.md1=lmer(AUTONOMY.LEVEL~RESIDENT.PGY.LEVEL..OP+(1|resid.hand), data=Hdata)
# sum.LMM.md1=summary(LMM.md1)
# kable(sum.LMM.md1$coefficients)

# LMM.md2=lmer(AUTONOMY.LEVEL~Type+RESIDENT.PGY.LEVEL..OP+(1|resid.hand), data=Hdata)
# summary(LMM.md2)
# With type as numeric
# LMM.num2=lmer(AUTONOMY.LEVEL~coded.Type+RESIDENT.PGY.LEVEL..OP+(1|resid.hand), data=
# sum.LMM.num2=summary(LMM.num2)
# kable(sum.LMM.num2$coefficients, caption = "Table of Estimates")
# read the LLM model with R examples
# LMM.null=lmer(AUTONOMY.LEVEL~RESIDENT.PGY.LEVEL..OP+(1|resid.hand), data=Hdata, REML

# full model
# LMM.full=lmer(AUTONOMY.LEVEL~Type+RESIDENT.PGY.LEVEL..OP+(1|resid.hand), data=Hdata,
```