Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Tianyi XU Date

Impacts and Improvement in Delineation of the Lumpectomy Cavity Boost Volume by using 3-D Implantable Tissue Markers

By

Tianyi Xu

Master of Science in Public Health

Biostatistics and Bioinformatics

Zhengjia (Nelson) Chen, PhD

(Thesis Advisor)

Azhar Nizam, MS

(Reader)

Impacts and Improvement in Delineation of the Lumpectomy Cavity Boost Volume by using 3-D Implantable Tissue Markers

By

Tianyi Xu

B.A.

Sun Yat-sen University

2016

Thesis Committee Chair: Zhengjia (Nelson) Chen, PhD

Reader: Azhar Nizam, MS

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health,
in Biostatistics and Bioinformatics
2018

Abstract

Impacts and Improvement in Delineation of the Lumpectomy Cavity Boost Volume by using 3-D Implantable Tissue Markers

By Tianyi Xu

Background: For patients undergoing breast conserving surgery (BCS), adjuvant radiation therapy (RT) is the standard of care. For patients with high risk features lumpectomy cavity boost (LCB) is typically added in. There can be much discrepancy in defining these cavities when guided by the traditional methods. The ambiguity may lead to either overestimate or underestimate of the cavity's actual size and affect the outcome, especially in the case of delay between surgery and adjuvant RT and oncoplastic reduction (OR). The dose RT boost Volume will be determined by the delineation of the cavity volume.

Methods and Materials: The records of 20 consecutive patients (41 to 76 years) who underwent BCS followed by adjuvant RT with a LCB at our institution were reviewed. Eight patients underwent BCS alone (40%) and 12 underwent BCS and OR (60%). At the time of surgery all patients had been placed a 3-D TM. The LCB volumes are determined by ten independent radiation oncologists. Paired T-tests and GEE model were used for statistical analysis.

Results: The mean LCB volume by using traditional methods was 33.2 cc (SD=29.2, SEM=6.9), while by using 3-D TM, the mean LCB volume was 13.2 cc (SD=13.8, SEM=2.5, P-value=0.072). By using the traditional methods, there was an average volume difference of being 17.4 cc larger than the actual treated boost volume. The corresponding percent overlap is 52% and an average DICE coefficient is 0.58.By using the 3-D TM method, the average volume difference was 2.6 cc smaller than the actual treated boost volume per patient. The corresponding percent overlap is 83.9% and an average DICE coefficient is 0.66.

Conclusion: Using 3-D TM method for planning will improve LCB volumes' reproducibility and accuracy. On average, it gains a smaller and more precise LCB volumes than traditional methods. Moreover, using the 3-D TM can improve inter-rater reliability, which is consistent in different levels of raters. Higher level raters perform better than lower level raters, and it is more obvious by using 3-D TM method. Thus using novel 3-D implantable tissue markers is better in improving consistency among radiation oncologist boost volume contours.

Impacts and Improvement in Delineation of the Lumpectomy Cavity Boost Volume by using 3-D Implantable Tissue Markers

By

Tianyi Xu

B.S.

Sun Yat-sen University

2016

Thesis Committee Chair: Zhengjia (Nelson) Chen, PhD

Reader: Azhar Nizam, MS

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health,
Biostatistics and Bioinformatics
2018

ACKNOWLEDGEMENT

I would first like to thank my thesis advisor Prof. Nelson Chen of Rollins School of Public Health at Emory University. He gave me a great help in academic field, answer my question patiently, taught me how to perform a proper research and keep steering me in the right direction. I would also like to thank my thesis reader Prof. Azhar Nizam for his great help during the revision process. He gave me detailed revision advice to make my thesis more reasonable. Finally, I must express my very profound gratitude to my parents for providing me with unfailing support and continuous encouragement throughout my years of study. Thank you.

CONTENTS

1. INTRODUCTION

In recent decades, the surgical management of breast cancer has steadily and considerably improved. Mutilating procedures have given way to more individualized surgical approaches aiming to preserve the breast as much as possible (Clough, 2015). Breast-conserving surgery (BCS, also known as segmental mastectomy) is a less radical cancer surgery than mastectomy. In a lumpectomy, breast-conserving surgery removes just part of the breast tissue as opposed to the entire breast. Adjuvant Radiation therapy is the standard of care for patients with breast cancer after breast-conserving therapy. For patients with high risk features (such as very young or old, positive margins, and high-grade disease) lumpectomy cavity boost (LCB) is often additionally performed. The process of defining the boost volume is important for the further care and recovery.

Traditionally, we use the post-operative seroma, anatomical landmarks, and surgical clips as a guidance to define the boost volume of LCB. The defining process may be dependent on the patients' age, body habitus and differentiation between breast tissue and fatty tissue (Pitkanen, 2001). Thus by using those conventional methods, there can inevitably be much discrepancy in defining these cavities. The ambiguity may lead to either overestimate or underestimate of the cavity's actual size and affect the outcome, especially in the case of delay between surgery and adjuvant RT and oncoplastic reduction (OR). If we can obtain an accurate delineation of the cavity volume, the RT boost volume will be more suitable for the patient, then we will reduce unnecessary side effects and achieve better therapeutic results.

In our study, there were several questions we were interested in. First, is there any improvement of cavity volume estimate and its preciseness by using 3-D TM methods compared to conventional methods? Second, does the 3-D methods affect the inter-observer reliability? After that, I used GEE model to determine relationships between corresponding variables, evaluating potential bias among measurement and rating process. The ultimate goal

is to evaluate the whether the use of a novel 3-D TM will improve reproducibility of the LCB volume, decrease ambiguity in the delineation of the surgical surgical cavity, and result in smaller, more accurate LCB volumes.

2. METHODS AND MATERIALS

The data for analysis are from a retrospective study on female patients performed by Winship Cancer Institute of Emory University, and the use of the data for this paper has been approved by the Emory University Institutional Review Board. Demographic data, such as age and gender, and related clinical features were obtained from the electronic medical records.

2.1 Patients and Observation parameter.

The dataset includes 20 consecutive female patients (41 to 76 years old, mean age is 61.55 years) who underwent breast-conserving surgery followed by adjuvant RT with a LCB at Winship Cancer Institute of Emory University between January 2015 and October 2017 were reviewed. Among these patients, eight of them underwent BCS alone (40%) and twelve underwent BCS and OR (60%). All patients had a 3-D TM placed at the time of surgery.

Demographic and physical indicators (age, sex and BMI), past medical history and past treatment history (Oncoplastic Reduction, Diagnosis, Stage, N-stage, T-stage, ER+, PR+, HER-2, Neo-adjuvant Chemo, Adjuvant Chemo, Chemo) were recorded to allow for control of covariates and quantification of the impact of past medical history and treatment history in the analysis of outcomes.

As for delineating the LCB volume, all the data were divided into two groups. (AC and BC) group. Within each group, every rater will give an individual estimation of those variables include the LCB Volume, Volume difference, Volume overlap, Percent overlap % and Dice

coefficient. Thus the whole dataset will have 400 observations, which is 20 patients * 2 group * 10 raters. Each observation's data will include those parameters estimation. By measuring and recording this, we can not only examine the reliability of the rater and the accuracy across different levels of resident physicians, but also decrease the bias during volume estimation.

2.2 Adjusting for Raters' accuracy

We invited ten independent resident physicians to evaluate each patients' LCB volumes. Those ten raters are resident physicians randomly selected from level 3 to 5 from Winship Cancer Institute of Emory University. Specifically, we coded the name of raters as the abbreviation of their name, in order to keep the privacy. The resident physician MA, ST and ZB are from grade 3; resident physicians BP, JP, NM and TM are from grade 4; resident physicians DZ, JJ and RC are from grade 5.

Besides estimating the parameters corresponding to LCB volume above, each rater need to finish a survey of self-estimation. The survey consists of those variables include Cavity Visualization Score, Level of Confidence in Boost Volume without Biozorb, Biozorb Visibility, Level of Confidence in Boost Volume with Biozrob, and Biozorb Utility for Boost or Primary Planning.

Specifically, the cavity visualization score (CVS) ranged from 1 to 5, to represent different levels of accuracy of delineation of LCB volume. We used 1 to represent no visible cavity; 2 for heterogeneous cavity with indistinct margins; 3 for heterogeneous cavity with distinct margins; 4 for mildly heterogeneous cavity with mostly distinct margins and 5 for homogenous cavity with clearly defined margins. The variable Level of Confidence in Boost Volume with/without Biozorb ranged from 1 to 5, where successively represents very confident, confident, neither confident nor unconfident, unconfident and very unconfident. The variable Biozorb Visibility ranged from 1 to 4, where successively represents easily visible, fairly visible, barely visible and not visible. The variable Biozorb Utility for Boost or

Primary Planning ranged from 1 to 4, and successively represents very helpful, helpful, unhelpful and very unhelpful.

Through this survey, we adjusted the impact of the raters during parameter estimation. For example, we can know how much confidence they have when the outline is not particularly clear to recognize. Also, we can discover the relationship between those variables and the Cavity Visualization Score, and adjust the potential bias during the experiment. After considering the impact of those variables, then we performed further research in the impacts of 3-D TM methods in delineation of the LCB volume, and whether the 3-D TM methods improve the reproducibility of the LCB volume.

2.3 Statistical Analysis

2.3.1. Descriptive analysis

The descriptive table for patients' characteristics and five LCB volume variables was firstly constructed. We hold the following criteria that, for continuous variables, the mean and standard deviation were summarized. For binary and categorical variables, the frequencies and percentage were presented. For five Volume Variables, we divided the whole dataset into two groups: the AC group and BC group (using the traditional indicators and the 3-D Implantable Tissue Markers). For each variable, we examined whether there exists interaction among different levels of the raters. So we calculated the mean within each PGY level separately to see the difference, and the standard deviation to compare the difference of accuracy.

2.3.2 Statistical analysis

First, we wanted to discover the difference between AC and BC group. Because AC and BC were measured within the same person, thus we have to use paired T-test for those observation parameters. Second, we wanted to know whether there exists any possible

interaction within grades of the residents (PGY level). So we did the same paired T-test among PGY 3, 4 and 5 separately. Third, we also want to examine the reliability of those raters, so the pairwise Pearson correlation coefficient (PCC) was calculated with the formula $\rho_{xy} = \frac{E[(X-\mu_x)(Y-\mu_y)]}{\sigma_x \sigma_y}$, where μ_x and μ_y are the mean of the two group, σ_x and σ_y are the standard deviation of the two group. PCC can measure the linear correlation between each rater and the expert, and it values from range from -1 to 1, with a value of 1 implying a perfectly linear relationship between variables. As for determining whether the rater's reliability while using the different methods (traditional methods or 3-D TM method), we performed this analysis separately in the subgroup of AC and BC. Also, we performed the analysis separately in the individual rater's level, thus we can see the difference between raters, and difference between levels. The 95% confidence interval were calculated for each estimate in order to show the significance.

After that, we wanted to know the accuracy rate of the raters' estimate, so the Dice Similarity coefficient (DSC) was estimated for each rater volume compared with the baseline volume using the formula $DSC = \frac{2*(A \cap B)}{A+B}$, where A is the baseline volume, B is the volume of a given rater, and $A \cap B$ is the intersecting overlap of the two volumes (Dice, 1945). A DSC > 0.7 is considered good overlap, with 1 being perfect overlap (Zijdenbos, 1994). Mean DSCs and SDs were calculate across each sample and each rater. As for convenience in comparison between different methods (AC/BC), different levels (PGY 3/4/5), and different raters (MA, ST, ZB, BP, JP, NM, TM, DZ, JJ and RC), we calculated DSCs separately in those subgroups. By examining the DSCs, we can know and compare the accuracy rate across different levels.

2.3.3 Univariate and Multivariate Regression Analysis.

In this study, the variables that we were interested in were almost all categorical variables. For example, there were 3 different levels of the raters (PGY 3, PGY 4 and PGY 5). As for

those variables with more than 4 categories, we split them from the middle and set them into 2 levels. For example, we set the Confidence Boost Volume level 1-2 as a lower level, and level 3-5 as a higher level. This setting will be more convenient for the data analysis and outcome explanation. At the very beginning, we did not know whether there exists possible mis-specified correlation structure. Thus the GEE model were chosen to fit the regression model, because it has consistent estimation even with mis-specified correlation structure. Also, the GEE model is computationally more simple for categorical data and does not require multivariate distribution (Hubbard, 2010). We fitted the data with a generalized estimating equation (GEE) model to test whether there was any significant change by PGY of each outcome and test whether there was any significant difference of each outcome between AC/BC. The model is shown here:

$$Model\ 1: \quad Y_{ij} = \beta_{0j} + \beta_{1j} * PGY_{ij} + \varepsilon_{ij},$$

Where $i = AC\ or\ BC\ method, j = PGY\ level\ from\ 3\ to\ 5$. Here the Y can be either Structure Volume, Volume Difference, Volume Overlap, Percent Overlap or Dice Coefficient. Also, we performed a two-sided t-test for the difference of beta coefficient between PGY level and reference group. So by this model setting, we can compare not only the method difference but also the difference of PGY level stratification. Then a multivariate GEE model were performed to test whether there was any significant impact between individual characteristics, raters' confidence during estimation and their CVS score output. The full model is shown here:

$$Model\ 2: \quad CVS_j = \beta_{0j} + \sum_{i=1}^{14} \beta_{ij} * x_{ij} + \varepsilon_j, \quad \epsilon_j \sim N(\mu_j, \sigma_j)$$

Here the x_i's were PGY, Stage, Diagnosis, Oncoplastic Reduction, Neo-adjuvant Chemo, Adjuvant Chemo, Chemo, Confidence Boost Volume with/without Biozorb, Biozorb Utility Boost, Biozorb Visibility, Age, BMI and Days after Radiation Surgery. Among those categorical variables, we chose one of the categories as reference group and presented the odds ratio of other groups' beta estimate, in order to see whether there was a significant difference between each categories. The j is representing Method levels, which is similar to

the model 1's setting. This model was by a backward variable selection method with an $\alpha = 0.02$ removal criteria. The significant level is set at $\alpha = 0.05$. The data in this study were analyzed using R version 3.2.3 and confirmed the outcome by SAS 9.4.

3. RESULTS

3.1 Results of Descriptive analysis.

The data structure of our study was shown by the **Figure 1 and 2**. The 400 observations of 20 patients were consisted of two methods groups (AC/BC), and 3 levels (PGY 3/4/5) of 10 raters. The descriptive statistics for patients' characteristics was shown by the **Table 1.a**. From the table, we can know the mean age of 20 female patients who received breast conserving surgery was 61.55 years old, and the mean BMI was 32.44. 60% patients had oncoplastic reduction after the surgery, 60% patients had invasive carcinoma, which is the most common type of breast cancer. (Adachi Y, 2018) The percentage of the stage of breast cancer 0, IA and IIA were 35%, 30% and 35%. As for pathologic stage of breast cancer, most of the patients have T-stage breast cancer were in Tis stage (35%). Most of the patients have N-stage breast cancer were in 0 stage (75%). As we know, cancerous cells of breast cancer may have none, one, or both receptors. Breast cancers that have estrogen receptors are called ER+, those with progesterone receptors are referred to as PR+, and in addition to hormone receptors, some breast cancers have high levels of a growth promoting protein called HER-2. (Carol A.Parise, 2014) In our study, 50% patients were diagnosed as ER-positive (or ER+), 40% patients were diagnosed as PR-positive (or PR+) and 10% patients were diagnosed as HER-2 negative breast cancer. As for chemotherapy history, 45% patients had received chemotherapy. More specifically, 20% patients had received neo-adjuvant chemotherapy while 5% patients had received adjuvant chemotherapy.

The results of descriptive analysis for LCB volume parameters were shown in **Table 1.b.** The mean Volume estimate in AC group is 33.23 while the mean Volume estimate in BC group is

13.18. More specifically, the mean Volume estimate by using AC method among PGY 3/4/5 are 32.83 (standard deviation=26.3), 36.87 (35.06) and 28.78 (22.5). The mean Volume estimate by using BC method among PGY 3/4/5 are 15.7 (standard deviation=18.75), 14.86 (13.26) and 8.40 (4.85). We can see that by using AC method, there is no trend of smaller Volume mean with increasing PGY level. While by using BC method, there is a trend of smaller Volume mean estimate with increasing PGY level. This outcome indicates that by using 3-D TM method, the raters are likely to conclude a smaller (accurate) estimation of LCB volume than traditional methods. The higher level the raters have, the better accuracy they will get. Moreover, when we look at the same PGY level, the BC group has a smaller standard deviation of the estimation, which means in in 3-D TM method, the estimation of 10 raters has smaller variance than traditional method.

We also noticed that the results above still exist among other variables, including Volume Difference, Volume Overlap, Percent Overlap and Dice coefficient. As for the Percent Overlap, we can see that in AC group among PGY 3/4/5, the estimate means are 52.78 (29.92), 51.19 (24.56) and 53.53 (27.59). While in BC group, those estimates are 81.70 (24.36), 80.06 (22.04) and 91.25 (13.15). Thus we can know that the BC group has a better overlap rate with the baseline volume. As for Dice Coefficient, we can see that in AC group among PGY 3/4/5, the estimate means are 0.57 (0.23), 0.6 (0.19) and 0.58 (0.21). While in BC group, those estimates are 0.65 (0.2), 0.68 (0.17) and 0.63 (0.2). Thus among all the PGY levels, using 3-D TM method will have a better similarity compared with the baseline estimate than traditional method. The outcome of significance test (a one-sided paired T-test) will be reported in the following part.

3.2 Results of Statistical analysis

As for **Table 1.b**, we preformed paired T-test for the difference of those 5 LCB Volume Parameters between AC and BC group. The null hypothesis is $H_0: AC < BC$, by comparing

corresponding p-value with the significance level $\alpha = 0.05$, we found that when using the BC method, Volume and Volume Overlap estimate are significantly smaller than using AC. As we can see from the Percent Overlap, although BC group's Volume estimate is smaller than AC, the BC's Percent Overlap is significantly lager than AC. This confirmed that raters' performance or accuracy is significantly better by using 3-D TM than traditional method, and the smaller Volume estimate indicates smaller blur area and more precise outline. Moreover, this results are still consistent when we examine at each PGY levels.

The Pearson Correlation Coefficient (PCC) and relative confidence interval were shown in **Table 2.** The Overall PCC in PGY 3/4/5 are 0.325 (0.155, 0.477), 0.240 (0.088, 0.381) and 0.391 (0.228, 0.533). To our surprise, the PGY 4 has the lowest PCC estimate, even lower than PGY 3. When we compared the AC and BC group, we found that by using AC method, PGY 4 has the lowest PCC estimate. But by using BC method, there still exists a trend of the higher PGY level, the higher PCC estimate. Thus we try to examine back to the AC group, then we found that the resident physician TM from PGY 4 has an extremely worse performance, which can be regarded as an outlier. His PCC in AC group is 0.006 (-0.438, 0.447) while the other resident physicians in PGY 4 has an average of 0.520 in AC group. So the extreme value might be the reason of the trend does not exist in AC group. When we exclude the TM's estimate, the trend also appears in AC group. Those results indicate that, the higher PGY level, the higher linear correlation between their estimate and baseline.

The Dice Similarity Coefficient (DSC) of rater's volume and baseline volume were shown in **Table 3.** For the overall comparison between AC and BC group, the overall mean of AC is 0.58 (0.141) while the overall mean of BC is 0.65 (0.141). Then we use a one-sided T-test to examine whether the BC is larger than AC mean. The null hypothesis is BC > AC with the corresponding p-value = 0.996. This results indicate that the 3-D TM group has a better Dice Similarity Coefficient than traditional group. We also noticed that 70% patients have an equal or larger DSC in BC group than AC. This indicates that compared to using traditional

method, the advantage of using 3-D TM method is obvious. When we consider the stratification in PGY level (among PGY 3/4/5), there are 80%, 70% and 65% patients have an equal or larger DSC in BC group than AC. So this confirms the advantage of using BC method, but the advantage is more obvious in lower level raters. This is a little bit counter-intuitive, because in the above analysis, we have found that 3-D TM might be a better method, and the higher level raters might perform a better estimate than lower level raters. In order to have a more thorough understanding of this difference between PGY levels, we repeated the calculation in each level separately, and presented the outcome in **Table 4.a**, **4.b** and **4.c**. We found that if we calculated the mean DSC separately of individual rater, then counted up to compare AC and BC group, among PGY 3/4/5, there were 73.3%, 57.5% and 61.6%. Thus we can see that PGY 3 and 5 has the similar estimate as before, while PGY 4 has decreased significantly from 70% to 57.5%. By examining the table, we found that several estimate are relatively close, for example, 0.62 is greater than 0.61. After exclude the paired data which their subtraction less than 0.02, we can found that the percentage became 69.3% which was close the 70%. Considering that the higher level raters have more accuracy, we calculate the standard deviation of those mean DSC estimate. The standard deviation of using AC method in PGY 3/4/5 are 0.226, 0.193 and 0.107; the standard deviation of using BC method in PGY 3/4/5 are 0.202, 0.165 and 0.101. Thus we realized that because higher level of rater has a more stable play (smaller standard deviation), thus their estimate in AC and BC are relatively close. So we can't directly compare the absolute value and conclude that the advantage of using BC method is more obvious in lower level raters.

3.3 Results of Univariate and Multivariate Regression Analysis

3.3.1 Results of univariate regression.

As for **Table 5**, we present the outcome of univariate association between PGY level and LCB Volume parameters, include Structure Volume, Volume Difference, Volume Overlap, Percent Overlap and Dice Coefficient. We typically focused on the analysis of Structure

Volume and Percent Overlap, because the Volume Difference is affected by the absolute value of Structure Volume and its meaning can be more properly described as Percent Overlap. We perform the model by using traditional method and 3-D TM method, and set PGY 5 as a reference group for the convenience of PGY- level comparison. The basic model is here,

$$Model\ 1: \quad Y_{ij} = \beta_{0j} + \beta_{1j} * PGY_{ij} + \varepsilon_{ij},$$

$$where\ i = AC\ or\ BC\ method, j = PGY\ level\ from\ 3\ to\ 5$$

First, we are coming to the analysis of Structure Volume. By using traditional method, the relative beta estimate of PGY 3 and 4 are 4.06 (-0.40, 8.51) and 8.10 (1.91, 14.28). And we performed a t-test of the difference between PGY levels and reference group. The corresponding p-value are 0.074 and 0.010. Refer to the conclusion from statistical analysis part, this outcome means that by using traditional method, there does not exist a trend that the higher PGY level, the smaller (or more accurate) Structure Volume estimate they will get. However, by using 3-D TM method, the relative beta estimate of PGY 3 and 4 are 7.30 (3.26, 11.34) and 6.47 (3.82, 9.11), and the corresponding p-value are both less than 0.01. This indicates that by using the 3-D TM method, when PGY level changes from 3 to 5, the relative Structure Volume is going to be smaller. Second, we are coming to the analysis of Percent Overlap. By using traditional method, the relative beta estimate of PGY 3 and 4 are -0.75 (-6.63, 5.12) and -2.34 (-6.64, 1.96), and the relative p-value compared to reference group are 0.802 and 0.286. By using 3-D TM method, the relative beta estimate of PGY 3 and 4 are -9.55 (-13.81, -5.28) and -11.19 (-15.04, -7.34), and the relative p-value compared to reference group are both less than 0.001. This outcome indicates that, by using the 3-D TM method, the difference between PGY level becomes more significant, and the highest level of rater will have a best performance in delineation of LCB volume. Due to some outlier (extreme value), the PGY 4 level has the worst overall performance.

3.3.2 Results of multivariate regression.

We constructed **Table 6** for the descriptive statistics of the variables in the multivariate regression, and constructed **Table 7** for the outcome of GEE model and relative p-value of multivariate regression. The basic model is here:

$$Model\ 2: \quad CVS_j = \beta_{0j} + \sum_{i=1}^{14} \beta_{ij} * x_{ij} + \varepsilon_j, \quad \epsilon_j \sim N(\mu_j, \sigma_j)$$

By using this model, we want to examine which variables among patient's medical history, treatment history and raters' confidence will have a significant impact of raters' CVS score. By **Table 7,** we found that only few variable has the relative p-value less than 0.05. More specifically, the p-value of level 1-2 Confidence Boost Volume is less than 0.001. Which means that, if the confidence level of the rater itself is relatively low, it will have a significant impact on their performance of output a CVS score. But the characteristics of patients, the medical and treatment history will not be a problem for those radiation oncologists' performance. This outcome confirms the stability and reliability of our design of using Cavity Visualization Score as a parameter of measuring the LCB volume.

4. CONCLUSION

According to our study, using 3-D TM method for planning will gain a smaller and more precise LCB volumes than traditional methods on average. The delineation of cavity has smaller volumes and a larger overlap rate (AC volume < BC volume, BC more precise) Moreover, according to our study, using the 3-D TM can improve inter-rater reliability. By using 3-D TM methods, raters' dice similarity is significantly better than traditional method. This results are consistent when we examine at different levels of raters. We also found that, higher level raters are more reliable than lower level raters. Because the higher levels of raters, the higher linear correlation between their estimates and baseline. This founding is more obvious by using 3-D TM methods. Thus we conclude that, the 3-D TM appeared to be

more effective and accurate at identifying and maintaining the cavity for boost determination. It's a better method as a guidance to the delineation of cavity, and better in improving consistency among radiation oncologist boost volume contours.

5. DISCUSSION

First, as for multivariate regression, there are so many possible model to choose. Why we choose the Generalized Equation Estimation model is because at the very beginning, we did not have any multivariate distribution assumption. Also, we didn't know that whether there exists possible mis-specified correlation structure. GEE is proper for possible unknown correlation between outcomes (Liang, 1986). The focus of the GEE is on estimating the average response over the population ("population-averaged" effects) rather than the regression parameters that would enable prediction of the effect of changing one or more covariates on a given individual (James, 2003). This is suitable for our study's design, and the explanation of the results will be easier than considering the case under several assumptions.

As for limitations, first, the sample size is relatively small, because we only have 20 patients. Thus by using the traditional and 3-D TM method on the same person, we doubled our observations. Also, the sample size and ability of raters can also be better, thus we can gain a more precise estimate and have a relatively small variance. In our study, due to some extreme value (outlier) and relatively small sample size of raters, the results are definitely being affected, although this affection might be within an acceptable range. According to our study, the overall trend of preferring the 3-D TM method is quite obvious in both statistical analysis and regression analysis, under ideal conditions we still need more sample size, in order to decrease the bias and obtain a more reliable results. Second, due to the limitation of space, more test of correlation structure is needed. The correlation structure might be varied a lot due to the estimation, the model and the variable that we choose to measure. Although the model

from GEE model is consistent even when the covariance structure is wrongly specified, I think we need to use other possible regression model to double-check the outcome.

6. REFERENCE

1. Clough, K. B., Benyahi, D., Nos, C., Charles, C., & Sarfati, I. (2015). Oncoplastic Surgery: Pushing the Limits of Breast-Conserving Surgery. *The Breast Journal, 21*(2), 140-146. doi:10.1111/tbj.12372

2. Koukourakis, G. (2009, October 04). Radiation therapy for early breast cancer. Retrieved March 28, 2018, from https://link.springer.com/article/10.1007/s12094-009-0410-2

3. Pitkänen, M. A., Holli, K. A., Ojala, A. T., & Laippala, P. (n.d.). Quality assurance in radiotherapy of breast cancer--variability in planning target volume delineation. Retrieved March 28, 2018, from https://www.ncbi.nlm.nih.gov/pubmed/11321661

4. Hubbard, A. E., Ahern, J., Fleischer, N. L., Laan, M. V., Lippman, S. A., Jewell, N., . . . Satariano, W. A. (2010). To GEE or Not to GEE. *Epidemiology, 21*(4), 467-474. doi:10.1097/ede.0b013e3181caeb90

5. Dice, L. R. (1945). Measures of the Amount of Ecologic Association Between Species. *Ecology, 26*(3), 297-302. doi:10.2307/1932409

6. Zijdenbos, A., Dawant, B., Margolin, R., & Palmer, A. (1994). Morphometric analysis of white matter lesions in MR images: Method and validation. *IEEE Transactions on Medical Imaging, 13*(4), 716-724. doi:10.1109/42.363096

7. Adachi, Y., Sawaki, M., Hattori, M., Yoshimura, A., Gondo, N., Kotani, H., . . . Iwata, H. (2018). Comparison of sentinel lymph node biopsy between invasive lobular carcinoma and invasive ductal carcinoma. *Breast Cancer*. doi:10.1007/s12282-018-0852-x

8. Parise, C. A., & Caggiano, V. (2014). Breast Cancer Survival Defined by the ER/PR/HER2 Subtypes and a Surrogate Classification according to Tumor Grade and Immunohistochemical Biomarkers. *Journal of Cancer Epidemiology, 2014*, 1-11. doi:10.1155/2014/469251

9. Kung-Yee Liang and Scott Zeger (1986). Longitudinal data analysis using generalized linear models. *Biometrika*. **73** (1): 13–22. doi*:*10.1093/biomet/73.1.13

10. Hardin, James; Hilbe Joseph (2003). Generalized Estimating Equations. London: Chapman and Hall/CRC. ISBN 1-58488-307-3.

7. FIGURES AND TABLES

Figure 1: Logical diagram of study design

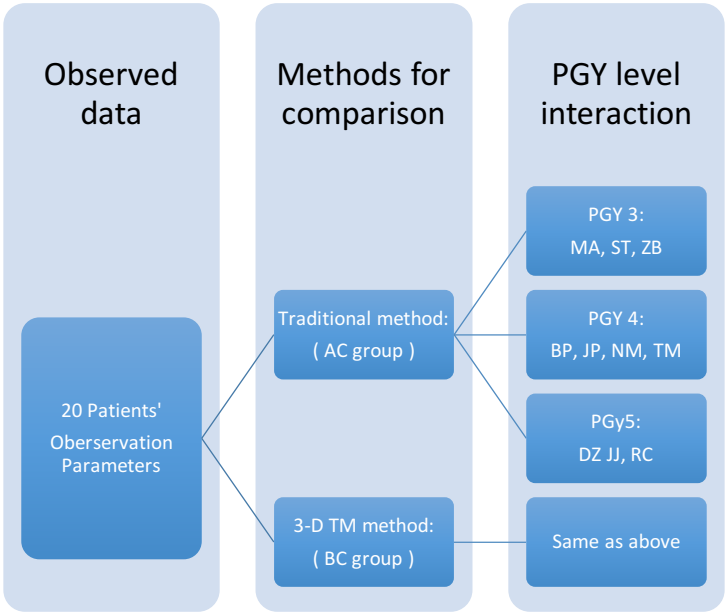

Figure 2. 20 Patient's volume estimate of different level of PGY

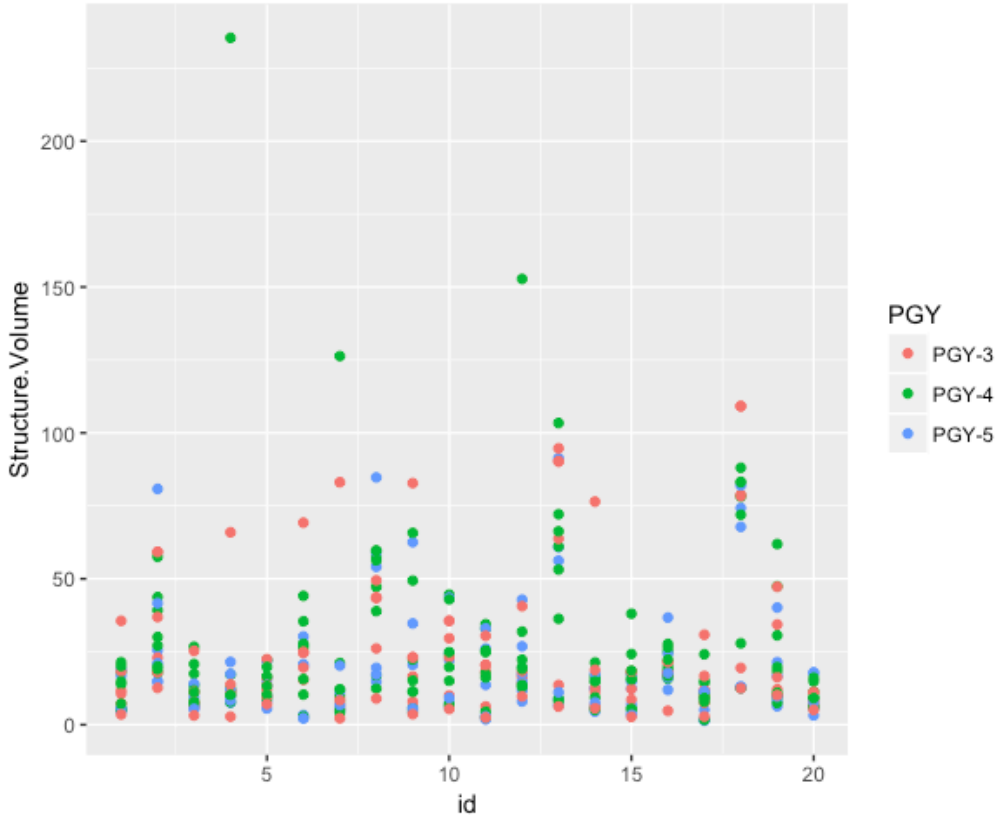

Table 1a: Descriptive Statistics of Patients' Characteristics Variables

Characteristics	N (%)	Characteristics	N (%)
Age, Years	n: 20	**ER+**	10 (50%)
from [41, 76]	mean: 61.55	No=0	10 (50%)
		Yes=1	
Sex, %		**PR+**	
Male	0 (0%)	No=0	12 (60%)
Female	20 (100%)	Yes=1	8 (40%)
Oncoplastic Reduction		**HER-2**	
Yes=1	12 (60%)	No=0	18 (90%)
No=2	8 (40%)	Yes=1	2 (10%)
Diagnosis		**Neo-adjuvant Chemo**	
0=DCIS	8 (40%)	No=0	16 (80%)
1=Invasive carcinoma	12 (60%)	Yes=1	4 (20%)
Stage		**Adjuvant Chemo**	
0=0	7 (35%)	No=0	15 (75%)
1=1A	6 (30%)	Yes=1	5 (25%)
2=IIA	7 (35%)		
T-stage		**Chemo**	
0=Tis	7 (35%)	No=0	11 (55%)
1=T1a	0 (0%)	Yes=1	9 (45%)
2=T1b	3 (15%)		
3=T1c	4 (20%)		
4=T2	6 (30%)		
N-stage		**BMI**	n: 20
0=0	15 (75%)	from [25.4, 50.7]	mean: 32.44
1=1a	1 (5%)		
2=Nx	4 (20%)		

Table1b: Descriptive Statistics of LCB Volume Parameters

Characteristics	**Total**	**AC (n=20) Mean(sd)**	**BC (n=20) Mean(sd)**	**P-value** H1: AC>BC Paired t-test
Volume	40	33.23	13.18	1.008e-06***
PGY-3	12	32.83(26.3)	15.70(18.8)	2.137e-06***
PGY-4	16	36.87(35.1)	14.86(13.3)	3.224e-06 ***
PGY-5	12	28.78(22.5)	8.40(4.9)	3.964e-05 ***
Volume difference	40	-17.43	2.62	1
PGY-3	12	-17.03(24.6)	0.10(17.7)	1
PGY-4	16	-21.07(34.2)	0.94(11.8)	1
PGY-5	12	-12.98(19.3)	7.4(6.7)	1
Volume overlap	40	12.91	9.40	6.975e-05 ***
PGY-3	12	12.29(7.2)	9.85(6.3)	0.0004448 ***
PGY-4	16	13.89(6.8)	10.31(6.0)	6.936e-06 ***
PGY-5	12	12.22(7.7)	7.72(4.5)	0.0004996 ***
Percent overlap, %	40	52.37	83.91	1
PGY-3	12	52.78(29.9)	81.70(24.4)	1
PGY-4	16	51.19(24.6)	80.06(22.0)	1
PGY-5	12	53.53(27.6)	91.25(13.2)	1
Dice coefficient	40	0.58	0.66	0.9608
PGY-3	12	0.57(0.2)	0.65(0.2)	0.9897
PGY-4	16	0.6(0.2)	0.68(0.2)	0.9802
PGY-5	12	0.58(0.2)	0.63(0.2)	0.8093

*** *Statistical significant, a=0.05*

Table 2. The PCC for Each Resident's Volume Estimate vs. Original Volume

Resident Doctor Level	**PCC in AC (95% CI) (N=20)**	**PCC in BC (95% CI) (N=20)**	**Overall PCC (95% CI)**
PGY-3	0.362 (0.119, 0.564)	0.339 (0.093, 0.546)	0.325 (0.155,0.477)
MA	0.571 (0.172, 0.809)	0.732 (0.429, 0.887)	0.500 (0.224, 0.702)
ST	0.313 (-0.150, 0.664)	0.486 (0.055, 0.764)	0.386 (0.085, 0.623)
ZB	0.342 (-0.118, 0.681)	0.454 (0.014, 0.716)	0.236 (-0.081, 0.510)
PGY-4	0.214 (-0.006, 0.414)	0.474 (0.283, 0.628)	0.240 (0.088, 0.381)
BP	0.381 (-0.074, 0.705)	0.507 (0.084, 0.776)	0.371 (0.068, 0.612)
JP	0.536 (0.122, 0.791)	0.585 (0.193, 0.816)	0.387 (0.087, 0.624)
NM	0.642 (0.279, 0.845)	0.623 (0.250, 0.835)	0.458 (0.170, 0.673)
TM	0.006 (-0.438, 0.447)	0.710 (0.391, 0.877)	0.149 (-0.170, 0.440)
PGY-5	0.554 (0.350, 0.709)	0.532 (0.321, 0.692)	0.391 (0.228, 0.533)
DZ	0.546 (0.137, 0.796)	0.661 (0.308, 0.854)	0.405 (0.107, 0.636)
JJ	0.579 (0.184, 0.813)	0.599 (0.213, 0.823)	0.409 (0.113, 0.640)
RC	0.572 (0.173, 0.810)	0.441 (-0.002, 0.739)	0.379 (0.076, 0.617)

Table 3. The Dice Similarity Coefficient of **PGY level** relative to ORG_AC, and the Mean for each AC and BC Methods of All Cases.

	PGY-3		**PGY-4**		**PGY-5**		**Mean(Sd)**	
id	**AC**	**BC**	**AC**	**BC**	**AC**	**BC**	**AC**	**BC**
1	0.77	0.78	0.79	0.82	0.78	0.73	0.78(0.01)	0.78(0.045)
2	0.28	0.54	0.31	0.59	0.25	0.69	0.28(0.03)	0.61(0.076)
3	0.54	0.64	0.50	0.60	0.29	0.58	0.44(0.134)	0.61(0.031)
4	0.50	0.61	0.70	0.78	0.73	0.63	0.64(0.125)	0.67(0.093)
5	0.73	0.70	0.68	0.73	0.54	0.72	0.65(0.098)	0.72(0.015)
6	0.64	0.70	0.61	0.82	0.58	0.83	0.61(0.03)	0.78(0.072)
7	0.51	0.66	0.54	0.68	0.48	0.67	0.51(0.03)	0.67(0.01)
8	0.48	0.57	0.52	0.59	0.52	0.58	0.51(0.023)	0.58(0.01)
9	0.53	0.66	0.63	0.83	0.54	0.84	0.57(0.055)	0.78(0.101)
10	0.59	0.68	0.67	0.55	0.73	0.16	0.66(0.070)	0.46(0.271)
11	0.48	0.47	0.53	0.51	0.66	0.40	0.57(0.093)	0.46(0.056)
12	0.49	0.62	0.38	0.63	0.43	0.73	0.43(0.055)	0.66(0.061)
13	0.68	0.44	0.63	0.48	0.60	0.20	0.64(0.040)	0.37(0.151)
14	0.35	0.73	0.56	0.77	0.51	0.76	0.47(0.110)	0.75(0.021)
15	0.52	0.77	0.53	0.67	0.56	0.68	0.54(0.021)	0.71(0.055)
16	0.69	0.67	0.69	0.64	0.60	0.64	0.66(0.052)	0.65(0.017)
17	0.47	0.63	0.46	0.71	0.48	0.82	0.47(0.01)	0.72(0.095)
18	0.61	0.63	0.86	0.57	0.79	0.45	0.75(0.129)	0.55(0.092)
19	0.75	0.86	0.65	0.86	0.64	0.83	0.68(0.061)	0.85(0.017)
20	0.77	0.60	0.83	0.77	0.79	0.67	0.80(0.031)	0.68(0.085)
mean	0.57	0.65	0.60	0.68	0.58	0.63	0.58(0.141)	0.65(0.141)

Table 4a. The Dice similarity coefficient of **each resident** relative to ORG_AC, and the mean values for each AC and BC methods of all cases (within PGY 3).

PGY3	**MA**		**ST**		**ZB**		**Mean(Sd)**	
id	**AC**	**BC**	**AC**	**BC**	**AC**	**BC**	**AC**	**BC**
1	0.86	0.81	0.75	0.85	0.68	0.68	0.76(0.091)	0.78(0.089)
2	0.35	0.78	0.26	0.26	0.25	0.58	0.29(0.055)	0.54(0.262)
3	0.67	0.78	0.70	0.70	0.26	0.45	0.54(0.246)	0.64(0.172)
4	0.73	0.66	0.00	0.77	0.77	0.41	0.50(0.433)	0.61(0.184)
5	0.72	0.73	0.72	0.72	0.76	0.66	0.73(0.023)	0.70(0.038)
6	0.77	0.83	0.49	0.49	0.68	0.80	0.64(0.148)	0.70(0.194)
7	0.81	0.72	0.02	0.90	0.70	0.36	0.51(0.428)	0.66(0.275)
8	0.53	0.73	0.42	0.42	0.51	0.55	0.48(0.057)	0.57(0.156)
9	0.79	0.81	0.00	0.80	0.81	0.36	0.53(0.462)	0.65(0.254)
10	0.70	0.57	0.68	0.76	0.39	0.71	0.59(0.173)	0.68(0.098)
11	0.62	0.78	0.27	0.00	0.55	0.65	0.48(0.185)	0.48(0.418)
12	0.59	0.72	0.44	0.44	0.44	0.69	0.49(0.087)	0.62(0.154)
13	0.68	0.43	0.73	0.73	0.63	0.17	0.68(0.045)	0.44(0.275)
14	0.69	0.92	0.04	0.69	0.31	0.58	0.35(0.327)	0.73(0.173)
15	0.75	0.78	0.41	0.85	0.39	0.70	0.52(0.202)	0.77(0.070)
16	0.72	0.86	0.72	0.72	0.65	0.44	0.70(0.040)	0.67(0.214)
17	0.56	0.77	0.40	0.40	0.46	0.74	0.47(0.081)	0.64(0.206)
18	0.72	0.50	0.38	0.96	0.75	0.42	0.61(0.202)	0.63(0.291)
19	0.78	0.90	0.55	0.90	0.92	0.77	0.75(0.187)	0.86(0.075)
20	0.76	0.73	0.80	0.80	0.75	0.28	0.77(0.026)	0.60(0.282)
mean	0.69	0.74	0.44	0.66	0.58	0.55	0.57(0.226)	0.65(0.202)

Table 4b. The Dice similarity coefficient of **each resident** relative to ORG_AC, and the mean values for each AC and BC methods of all cases (within PGY 4).

PGY4	BP		JP		NM		TM		Mean(Sd)	
id	**AC**	**BC**	**AC**	**BC**	**AC**	**BC**	**AC**	**BC**	**AC**	**BC**
1	0.76	0.81	0.76	0.78	0.82	0.84	0.82	0.84	0.78(0.035)	0.08(0.030)
2	0.21	0.39	0.35	0.75	0.36	0.69	0.36	0.69	0.31(0.084)	0.61(0.193)
3	0.40	0.65	0.55	0.53	0.74	0.62	0.74	0.62	0.56(0.170)	0.60(0.062)
4	0.58	0.80	0.77	0.79	0.78	0.71	0.78	0.71	0.71(0.113)	0.77(0.049)
5	0.59	0.74	0.74	0.71	0.76	0.70	0.76	0.70	0.70(0.093)	0.72(0.021)
6	0.65	0.76	0.74	0.84	0.60	0.84	0.60	0.84	0.66(0.071)	0.81(0.046)
7	0.51	0.72	0.70	0.59	0.80	0.63	0.80	0.63	0.67(0.147)	0.65(0.067)
8	0.48	0.70	0.56	0.55	0.51	0.56	0.51	0.56	0.52(0.040)	0.60(0.084)
9	0.75	0.83	0.84	0.76	0.84	0.88	0.84	0.88	0.81(0.052)	0.82(0.060)
10	0.62	0.72	0.72	0.17	0.64	0.57	0.64	0.57	0.66(0.053)	0.49(0.284)
11	0.50	0.48	0.33	0.22	0.61	0.67	0.61	0.67	0.48(0.141)	0.46(0.226)
12	0.35	0.50	0.49	0.82	0.45	0.82	0.45	0.82	0.43(0.072)	0.71(0.185)
13	0.58	0.65	0.59	0.20	0.67	0.33	0.67	0.33	0.61(0.049)	0.39(0.232)
14	0.57	0.63	0.57	0.78	0.60	0.77	0.60	0.77	0.58(0.017)	0.73(0.084)
15	0.36	0.42	0.73	0.71	0.58	0.76	0.58	0.76	0.56(0.186)	0.63(0.184)
16	0.66	0.66	0.70	0.56	0.75	0.69	0.75	0.69	0.70(0.045)	0.64(0.068)
17	0.33	0.74	0.54	0.78	0.61	0.82	0.61	0.82	0.49(0.146)	0.78(0.040)
18	0.83	0.62	0.90	0.45	0.84	0.40	0.84	0.40	0.86(0.038)	0.49(0.115)
19	0.65	0.82	0.88	0.89	0.87	0.87	0.87	0.87	0.80(0.130)	0.86(0.036)
20	0.78	0.70	0.84	0.80	0.84	0.72	0.84	0.72	0.82(0.035)	0.74(0.053)
mean	0.56	0.67	0.66	0.63	0.68	0.69	0.51	0.73	0.60(0.193)	0.68(0.165)

Table 4c. The Dice similarity coefficient of **each resident** relative to ORG_AC, and the mean values for each AC and BC methods of all cases (within PGY 5).

PGY5	**DZ**		**JJ**		**RC**		**Mean(SD)**	
id	**AC**	**BC**	**AC**	**BC**	**AC**	**BC**	**AC**	**BC**
1	0.71	0.58	0.86	0.75	0.76	0.85	0.78(0.076)	0.73(0.137)
2	0.31	0.61	0.18	0.70	0.28	0.75	026.(0.068)	0.69(0.071)
3	0.33	0.52	0.00	0.59	0.53	0.62	0.29(0.268)	0.58(0.051)
4	0.79	0.63	0.63	0.63	0.78	0.61	0.73(0.090)	0.62(0.012)
5	0.43	0.69	0.65	0.72	0.55	0.74	0.54(0.110)	0.72(0.025)
6	0.74	0.77	0.38	0.88	0.61	0.84	0.58(0.182)	0.83(0.056)
7	0.71	0.64	0.00	0.64	0.73	0.74	0.48(0.416)	0.67(0.058)
8	0.53	0.59	0.54	0.59	0.50	0.57	0.52(0.021)	0.58(0.012)
9	0.88	0.75	0.00	0.98	0.74	0.80	0.54(0.473)	0.84(0.121)
10	0.77	0.23	0.64	0.15	0.79	0.10	0.73(0.081)	0.16(0.066)
11	0.81	0.27	0.59	0.33	0.59	0.59	0.66(0.127)	0.40(0.170)
12	0.51	0.73	0.34	0.74	0.43	0.74	0.43(0.085)	0.74(0.006)
13	0.66	0.15	0.52	0.23	0.63	0.22	0.60(0.074)	0.20(0.044)
14	0.47	0.72	0.51	0.76	0.54	0.80	0.51(0.035)	0.76(0.040)
15	0.55	0.74	0.66	0.59	0.47	0.71	0.56(0.095)	0.68(0.079)
16	0.65	0.60	0.63	0.69	0.52	0.64	0.60(0.070)	0.64(0.045)
17	0.57	0.80	0.53	0.86	0.34	0.80	0.48(0.123)	0.82(0.035)
18	0.76	0.37	0.77	0.45	0.84	0.55	0.79(0.044)	0.46(0.090)
19	0.71	0.68	0.47	0.90	0.73	0.90	0.64(0.145)	0.83(0.127)
20	0.74	0.58	0.81	0.69	0.82	0.75	0.79(0.044)	0.67(0.086)
mean	0.63	0.58	0.49	0.64	0.61	0.67	0.58(0.107)	0.63(0.101)

Table 5. Univariate Association (PGY)

Variable	***Covariate: PGY***		
Level	***3***	***4***	***5***
N-count	60	80	60
Structure Volume			
$\beta_{1(AC)}$	4.06 (-0.40, 8.51)	8.10 (1.91, 14.28)	Reference
P-value	0.074	0.010 ***	group
$\beta_{1(BC)}$	7.30 (3.26, 11.34)	6.47 (3.82, 9.11)	
P-value	< 0.01 ***	< 0.01 ***	
Volume Difference			
$\beta_{1(AC)}$	-4.06 (-8.51, 0.40)	8.10 (-14.28, 1.91)	Reference
P-value	0.074	0.010 ***	group
$\beta_{1(BC)}$	-7.30 (-11.34, -3.26)	-6.47 (-9.11, -3.82)	
P-value	< 0.01 ***	< 0.01 ***	
Volume Overlap			
$\beta_{1(AC)}$	0.07 (-0.63, 0.76)	1.67 (1.13, 2.20)	Reference
P-value	0.854	< 0.001 ***	group
$\beta_{1(BC)}$	2.13 (0.76, 3.50)	2.59 (1.58, 3.61)	
P-value	0.002 ***	< 0.001 ***	
Percent Overlap			
$\beta_{1(AC)}$	-0.75 (-6.63, 5.12)	-2.34 (-6.64, 1.96)	Reference
P-value	0.802	0.286	group
$\beta_{1(BC)}$	-9.55 (-13.81, -5.28)	-11.19 (-15.04, -7.34)	
P-value	< 0.001 ***	< 0.001 ***	
Dice Coefficient			
$\beta_{1(AC)}$	-0.01 (-0.06, 0.05)	0.03 (-0.004, 0.06)	Reference
P-value	0.852	0.083	group
$\beta_{1(BC)}$	0.02 (-0.05, 0.09)	0.05 (-0.001, 0.10)	
P-value	0.605	0.061	

**** Statistical significant, a=0.05*

Table 6. Descriptive Statistics (CVS)

	Cavity Visualization Score		
Covariate	***N-count***	***Level 1~2 N=79***	***Level 3~5, N=121***
PGY			
3	60	22 (36.67)	38 (63.33)
4	80	31 (38.75)	49 (61.25)
5	60	26 (43.33)	34 (56.67)
Stage			
0	70	29 (41.43)	41 (58.57)
1	60	30 (50)	30 (50)
2	70	20 (28.57)	50 (71.43)
Diagnosis			
0	80	35 (43.75)	45 (56.25)
1	120	44 (36.67)	76 (63.33)
Oncoplastic Reduction			
1=Yes	120	52 (43.33)	68 (56.67)
2=No	80	27 (33.75)	53 (66.25)
Neo-adjuvant Chemo			
0=No	160	64 (40)	96 (60)
1=Yes	40	15 (37.5)	25 (62.5)
Adjuvant Chemo			
0=No	150	63 (42)	87 (58)
1=Yes	50	16 (32)	34 (68)
Chemo			
0=No	110	48 (43.64)	62 (56.36)
1=Yes	90	31 (34.44)	59 (65.56)
Confidence Boost Volume			
Level 1-2	98	10 (10.2)	88 (89.8)
Level 3-5	102	69 (67.65)	33 (32.35)
Confidence BV with Biozorb			
Level 1-2	166	60 (36.14)	106 (63.86)
Level 3-5	34	19 (55.88)	15 (44.12)
Biozorb Utility Boost			
Level 1-2	147	65 (44.22)	82 (55.78)
Level 3-5	53	14 (26.42)	39 (73.58)
Biozorb Visibility			
Level 1-2	175	64 (36.57)	111 (63.43)
Level 3-5	25	15 (60)	10 (40)
Age	200	59.9 (9.59)	62.63 (9.59)
BMI	200	31.19 (5.01)	33.25 (7.41)
Days after Radiation Surgery	200	80.57 (45.52)	90.54 (54.9)

Table 7. Multivariate Association (CVS)

	Cavity Visualization Score		
Covariate	***N-count***	***Odds Ratio (95% CI)***	***P-value***
PGY			
3	60	1.32 (0.58, 3.03)	0.512
4	80	1.21 (0.76, 1.93)	0.430
5	60	Reference Group	
Stage			
0	70	0.57 (0.23, 1.41)	0.224
1	60	0.40 (0.13, 1.27)	0.120
2	70	Reference Group	
Diagnosis			
0	80	0.74 (0.34, 1.63)	0.461
1	120	Reference Group	
Oncoplastic Reduction			
1=Yes	120	0.67 (0.26, 1.70)	0.395
2=No	80	Reference Group	
Neo-adjuvant Chemo			
0=No	160	0.90 (0.31, 2.61)	0.846
1=Yes	40	Reference Group	
Adjuvant Chemo			
0=No	150	0.65 (0.21, 2.06)	0.462
1=Yes	50	Reference Group	
Chemo			
0=No	110	0.68 (0.28, 1.63)	0.386
1=Yes	90	Reference Group	
Confidence Boost Volume			
Level 1-2	98	16.34 (6.75, 38.87)	< 0.001 ***
Level 3-5	102	Reference Group	
Confidence BV with Biozorb			
Level 1-2	166	1.73 (0.76, 3.92)	0.192
Level 3-5	34	Reference Group	
Biozorb Utility Boost			
Level 1-2	147	0.41 (0.19, 0.89)	0.023
Level 3-5	53	Reference Group	
Biozorb Visibility			
Level 1-2	175	2.02 (0.71, 5.71)	0.186
Level 3-5	25	Reference Group	
Age	200	1.03 (1.00, 1.06)	0.074
BMI	200	1.05 (0.98, 1.12)	0.103
Days after Radiation Surgery	200	1.00 (1.00, 1.01)	0.351

**** Statistical significant, a=0.01*