BAYESIAN STOCHASTIC FRONTIER ANALYSIS WITH MISSING DATA MANAGEMENT AS KNOWLEDGE-BASED PLANNING FOR LUNG SBRT

Angelika Kroshko, Olivier Morin[1], Louis Archambault
CHU de Québec - Université Laval, Québec, Qc, Canada ; [1]University of California San Francisco, San Francisco, CA, USA

INTRODUCTION

Most knowledge-based planning (KBP) strategies rely on large volume of contoured data, but in clinical practice, only nearby OARs are delineated. To keep KBP clinically relevant, lengthy contouring sessions must be avoided. Our objectives are (1) to establish a novel based on multiple imputation to incorporate the data for patients with missing contoured structures and (2) to develop a knowledge-based planning method to predict optimal dosimetric indices for lung SBRT.

A Bayesian approach allows us to assess the influence of the parameters used for the variable selection and model specification. Stochastic frontier analysis allows to incorporate the effect of technical inefficiency in the model and to obtain a frontier optimizing the dose parameters (minimize them for the OARs or maximize them for the PTV) instead of a regression [1]. Missing data imputation provides a way to use all the available patient information and comply with fluctuating positions of the tumor in the lungs resulting in a variable dosimetric impact of the neighboring OARs in treatment planification.

METHODS

A retrospective study of 249 patients (219 training + 30 validation) treated for lung SBRT with VMAT technique was made. Prescribed dose to the PTV were 48 or 52 Gy in 4 fractions or 50 Gy in 5 fractions. Bayesian Stochastic Frontier Analysis (BSFA) is used to predict dosimetric indices based on anatomical features with a Markov Chain Monte Carlo algorithm. Geometric parameters were extracted between the PTV and 12 OARs such as, spinal cord, great vessels, esophagus, heart, main bronchus. Minimum distances between the PTV and the OARs for missing structures are estimated as following a truncated normal distribution with an established censoring value for its mean specific to each OARs (see table 1).

Geometric parameters are extracted with an automated script in RayStation. They are the minimum distance between the PTV and the OARs, the overlap volume between them if there is one, and the overlap volume between the OAR and a lateral and transversal expansion of PTV in order to evaluate the volume of the PTV that is in the field of treatment. Further planification parameters such as the prescribed dose are used in the predictive models.

Table 1: For the missing data imputation, a censoring value c_{OAR} for the minimum distance is chosen for every OARs. The standard deviation sd_{OAR} of the distribution of this parameter is also used. Then, we generate a value for the missing minimum distance following a truncated normal with its mean and lower bound being the censoring value. Variations of the imputation model are tested to determine if the method is robust.

Model	Distribution	μ	σ
A	$N^{+}(\mu,\sigma^2)$	c_{OAR}	sd_{OAR}
B	$N^{+}(\mu,\sigma^2)$	c_{OAR}	$2sd_{OAR}$
C	$N^{+}(\mu,\sigma^2)$	$c_{OAR} - 0.5$ cm	sd_{OAR}
D	$N^{+}(\mu,\sigma^2)$	$c_{OAR} + 0.5$ cm	sd_{OAR}
E	$N(\mu,\sigma^2)$	μ_{OAR}	sd_{OAR}

RESULTS

For the validation set, mean difference between observed and predicted values are -3.2 ± 2.3 % for V100% for the PTV, 2.4 ± 1.7 Gy for the D0.35cc and 2.1 ± 1.7 Gy for the Dmax of the spinal cord. Figure 1 presents the histogram of the difference between observed and predicted dose for the D0.35cc of the spinal cord which we try to minimize. Figure 2 presents the histogram for the V100% of the PTV which we try to maximize.

No significative changes in the parameter estimates between modification in the censoring values and the standard deviation of the distribution demonstrate that the implemented treatment of missing values is robust. Figure 3 presents the PTV V100% Bayesian Stochastic Frontier Analysis (BSFA) model estimate for the minimum distance between the PTV and the main bronchus for the five different missing data imputation models.

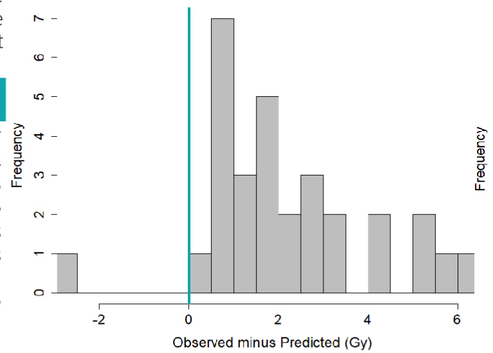

Figure 1. Histogram of the difference between observed and predicted value for the D0.35cc to the spinal cord for the validation cohort.

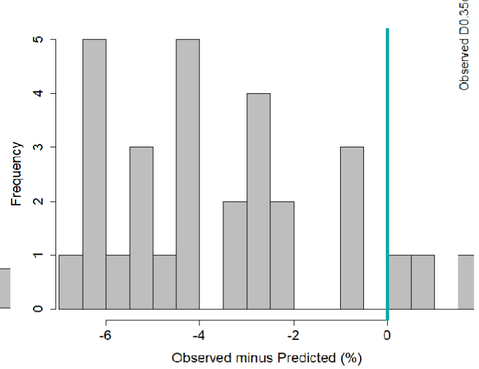

Figure 2. Histogram of the difference between observed and predicted value for the V100% to the PTV for the validation cohort.

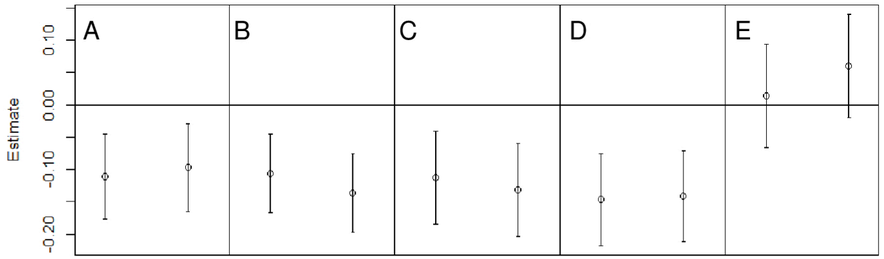

Figure 3. V100% for the PTV BSFA model estimate for the minimum distance between the PTV and main bronchus. Each missing data imputation model is repeated two times.

Furthermore, higher order effects of the minimum distance between PTV and spinal cord were significant parameters to predict the dose for this OAR. Figure 3 shows the effect of taking in account the higher order effect of the minimum distance to the spinal cord into the predictive model for the training set.

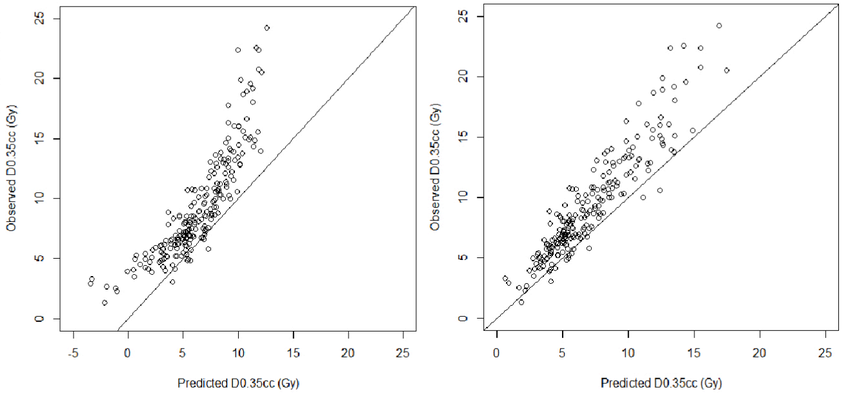

Figure 4. Predicted vs observed values of the training set for the D0.35cc to the spinal cord when higher order effect of the minimum distance between the PTV and the spinal cord aren't (left) and are (right) taken in account.

CONCLUSION

BSFA combined with our implemented missing data imputation proves to be a promising method to predict dosimetric parameters for PTV and OARs for lung SBRT. This approach allows to exploit larger database with missing contoured structures within a knowledge-based framework. This aligns to the reality of the clinical practice for complex radiation therapy techniques such as SBRT.

REFERENCES

[1] Angelika Kroshko, Olivier Morin, Louis Archambault. Stochastic frontier analysis as knowledge-based model to improve sparing of organs-at-risk for VMAT-treated prostate cancer. Physics in Medicine & Biology, Volume 64, no. 8 (2019)