First Latin-American Congress
on Solid State Dosimetry and Radiation Measurements
13-17 September, 2021

# AUTOMATIC WORKFLOW FOR GAMMA ANALYSES QA VERIFICATION

LIZAR, J.C.[1]; YALY, C.C.[2]; COLELLO. B.A.[2]; ARRUDA, G.V.[2]; PAVONI, J.F.[1,2]

[1]Department of Physics, Faculty of Philosophy, Sciences and Letters at Ribeirão Preto, University of São Paulo. Av. Bandeirantes 3900, 14040-901, Monte Alegre, Ribeirão Preto, São Paulo, Brazil, jclizar@usp.br
[2]Radiotherapy Department, Ribeirão Preto Medical School Hospital and Clinics, University of São Paulo. Av. Bandeirantes 3900, 14040-900, Monte Alegre, Ribeirão Preto, São Paulo, Brazil

**Introduction:** The gamma function is the standard methodology for dose comparation. This procedure is calculated on dedicated software, and no results verification is performed. Considering it, we developed a machine learning (ML) tool for patient-specific QA results verification. First, we evaluated the test gamma approval achieved with perpendicular composite dose distribution measurements based on gamma radiomics features [1]. Then, we analyzed the gamma process by verifying if the recommended software protocol was followed during the dose distribution comparison to detect possible users' errors.

**Material and method:** This study used 158 patient-specific IMRT QA tests, from different sites, and extracted 105 radiomic features from each gamma image. The analyses were done using Matrixx detector and OmniPro software (Iba Dosimetry). From this data, two datasets were randomly partitionated: dataset A was used to build the ML model, and dataset B was used as additional data to evaluate its performance. Three random forest (RF) models were developed (ML I, ML II, and ML III). ML I and ML II verified the features related to gamma image approval using criteria of 2%/2mm/15% threshold and 3%/3mm/15% threshold, respectively. ML III verified if the gamma analyses (3%/3mm/15% threshold) software recommended protocol was followed to detect if the TPS grid modification step was done. All models were based on the most important features selected using the mean decreased impurity, and their performances were evaluated. Approved plans received a class value = 1, and reproved ones a class value = 0.

**Results**: ML I included 25 features (5 First-order, 6 GLCM, 3 GLDM, 7 GLRLM, and 4 GLSZM), and its accuracy with the test set data was 0.98. The same value was achieved using the never-seen data (dataset B). ML II included 10 features (3 First-order, 3 GLCM, 1 GLDM, 1 GLRLM, and 2 GLZM.). Its accuracy was 0.85 using the test set data and 0.84 using dataset B. The First-order 10th percentile feature was identified as a feature strongly related to the approved classification for both models. ML III selected 23 features (9 GLCM, 5 GLDM, 3 GLRLM, 5 GLSZM, and 1 of NGTDM) with an accuracy of 0.99 for test set data and 0.98 for dataset B (Table 1).

**Conclusions:** An ML workflow for gamma analyses QA results verification could be proposed for 3%/3mm/95% criteria with the best results. We could verify the analysis protocol by checking the grid conversion during the gamma analyses (ML III). If the image pass in this classification, then we verify the approval or not of the QA (ML II) to see if its results behavior is according to expected. This procedure could cover an existing gap in the clinical practice of not verifying the gamma software algorithm's performance, such as done for the TPS algorithm.

Table 1: Evaluation parameters using the test set from dataset A and the additional dataset B.

| Models | Data | True label | F1-score | Accuracy |
|---|---|---|---|---|
| ML I | Test set | 0 | 0.87 | 0.85 |
| | | 1 | 0.83 | |
| | Dataset B | 0 | 0.87 | 0.84 |
| | | 1 | 0.77 | |
| ML II | Test set | 0 | 0.99 | 0.98 |
| | | 1 | 0.98 | |
| | Dataset B | 0 | 0.91 | 0.98 |
| | | 1 | 0.99 | |
| ML III | Test set | 0 | 0.99 | 0.99 |
| | | 1 | 0.99 | |
| | Dataset B | 0 | 0.98 | 0.98 |
| | | 1 | 0.98 | |

**References:**

1. van Griethuysen, J. J. M. et al. (2017). Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Research, 77(21), e104–e107. `