

# Dosimetry Study of Radiotherapy based on Data Mining

<sup>1</sup>Renwei JIANG, <sup>2</sup>Shuanshuan GUO, <sup>3</sup>Shuai Yang and <sup>4</sup>Qie DING,  
<sup>1,2,3,4</sup>Department of Radiation Oncology, The Fifth Affiliated Hospital, Sun Yat-sen University, Zhuhai, China

**Abstract:** Since the 1970s, in order to protect the psychological and physiological health of breast cancer patients and improve their quality of life, the scope of surgical treatment of breast cancer tends to become smaller. At present, breast-conserving surgery plus sentinel lymph node biopsy has gradually replaced radical surgery as the standard procedure for patients with early breast cancer. Early breast cancer patients after breast-conserving surgery require postoperative adjuvant radiation therapy. In recent years, conformal intensity-modulated radiation therapy has gradually replaced tangential field irradiation as the main mode of adjuvant radiation therapy after breast-conserving surgery. Conformal intensity-modulated radiation therapy has better target area conformity and dose uniformity in the target area than ordinary tangential field radiation therapy, and can better protect important dangerous organs. Compared with conventional forward radiotherapy techniques, the advantages of intensity-modulated radiotherapy in reducing pelvic dose have been well studied, but in IMRT, the number of fields is different, and the dosimetric protection of pelvis and normal tissues is different.

**Keywords:** Data mining, Radiotherapy technology, Dosimetry, Data acquisition, Data analysis

## I. INTRODUCTION

The advantage of three-dimensional conformal radiotherapy (3DCRT) is that it has an obvious dose distribution advantage, improves the target dose distribution, and reduces the exposure dose of normal tissue around the target area. Studies have shown that blood toxicity above 2 degrees is strongly associated with pelvic V20. Patients with V20 greater than 80% have a blood toxicity of more than 2 degrees 4.5 times that of patients with V20 less than 80%. Studies have shown that blood toxicity above 3 degrees is related to the average dose of pelvis, while the tibia, lower pelvis and clinical factors are irrelevant. At present, some researchers focus on the use of intensity-modulated radiation therapy (IMRT) to reduce the radiation dose of pelvic bone. Compared with conventional forward radiation therapy, the advantage of intensity-modulated radiation therapy in reducing the pelvic dose has been fully Studies, but in IMRT, the number of fields is different, and the dosimetric protection of pelvis and normal tissues is also different.

Local advanced non-small cell lung cancer refers to an important structure that has been associated with mediastinal lymph node and supraclavicular lymph node metastasis, invading the apex of the lung and mediastinum, and has not found distant metastasis of non-small cell lung cancer using existing examination methods. Important structures include pericardium, heart, large blood vessels, and esophagus. Among lung cancer cases, NSCLC accounts for more than 2/3, and most cases of stage IIIa and IIIb at the time of treatment are difficult to completely remove by surgery. The short-term and long-term effects of conventional radiotherapy for inoperable NSCLC patients are not optimistic.

Whole brain spinal cord irradiation is mainly used for the treatment of brain malignant tumors that are easily spread along the meninges or cerebrospinal fluid, and has a high 5-year survival rate. Since the CSI target region runs through the entire human brain to the concentric region of the conus of the spinal cord, the irradiation range is large, and patients are prone to major adverse reactions such as nausea, vomiting, and myelo suppression. Therefore, on the basis of ensuring that the whole central target area receives sufficient dose of radiation, how to further reduce the exposure dose of each dangerous organ becomes the focus of clinical research.

Rotary tomography and volumetric rotational intensity modulated radiation therapy have been widely used in clinical practice as the more advanced radiation therapy technology. HT performs spiral irradiation on complex target areas from 360° in 51 fields, with a maximum treatment length of up to 160 cm. It can maximize the protection of surrounding normal tissues while improving the conformality of the target area, and has the technical advantage of implementing CSI. VMAT is a dynamic intensity-modulation technique that achieves dose coverage of the target area by changing the dose rate and multi-leaf collimator field shape while rotating the gantry, and can achieve longer target areas through multi-center planning optimization. The highly conformal illumination makes it possible for clinical applications of total central illumination.

## II. THE PROPOSED METHODOLOGY

**Data Mining:** Considering the data itself, data mining usually requires eight steps: data cleansing, data transformation, data mining implementation process, pattern evaluation and knowledge representation. Information collection: Abstract the characteristic information needed in the data analysis according to the determined data analysis object, and then select the appropriate information collection method to store the collected information in the database. For massive amounts of data, choosing a data warehouse for proper data storage and management is critical. Data integration: The data of different sources, formats and characteristics are organically concentrated in a logical or physical manner, thus providing enterprises with comprehensive data sharing. Data Protocol: Performing most data mining algorithms takes a long time even on a small amount of data, while data mining in commercial operations often has a very large amount of data. The data specification technique can be used to obtain a protocol representation of the data set, which is much smaller, but still close to maintaining the integrity of the original data, and the data mining results after the specification are the same or nearly the same as the pre-statute execution results. Data cleansing: Some of the data in the database is incomplete (some attributes of interest are missing attribute values), noisy (including incorrect attribute values), and are inconsistent (the same information is represented differently). Therefore, data cleaning is required to store complete, correct, and consistent data information in the data warehouse. Data transformation: Convert data into a form suitable for data mining through smooth aggregation, data generalization, normalization, and so

on. For some real-type data, transforming data through concept stratification and data discretization is also an important step. Data mining process: According to the data information in the data warehouse, select the appropriate analysis tools, apply statistical methods, case reasoning, decision trees, rule reasoning, fuzzy sets, even neural networks, genetic algorithms to process information, and obtain useful analysis information. Model evaluation: From a business perspective, industry experts verify the correctness of data mining results. Knowledge representation: The analysis information obtained by data mining is presented to the user in a visual way, or stored as a new knowledge in the knowledge base for use by other applications.

Decision trees are an algorithm commonly used in predictive models. They find valuable and potential information by purposely classifying large amounts of data. Its main advantages are simple description and fast classification, which is especially suitable for large-scale data processing.

Neural network is very suitable for solving data mining problems due to its good robustness, self-organization self-adaptation, parallel processing, distributed storage and high fault tolerance. It is a feedforward neural network model for classification, prediction and pattern recognition.

Rough set theory is a mathematical tool for studying inaccurate and uncertain knowledge. The rough set method has several advantages: no need to give additional information; simplify the expression space of the input information; the algorithm is simple and easy to operate. The object of rough set processing is an information table similar to a two-dimensional relation table. But the mathematical basis of the rough set is set theory, it is difficult to directly deal with continuous attributes. The continuous attributes in the reality information table are ubiquitous. Therefore, the discretization of continuous attributes is a difficult point to restrict the practical application of rough set theory.

According to the information storage format, the objects used for mining include a relational database, an object-oriented database, a data warehouse, a text data source, a multimedia database, a spatial database, a temporal database, a heterogeneous database, and an internet.

#### **Dosimetric Comparison of Three-dimensional Conformal Radiotherapy and Intensity-modulated Radiotherapy:**

Firstly, the position is fixed, and the positioning bag is placed in the positioning frame. The patient's hands are lifted and placed on the top of the forehead, lying on the top of the positioning bag. After fully contacting, the positioning bag is vacuumed and sealed, and the laser positioning device is used in the human body and the frame. And the positioning bag is marked to make the body position consistent with CT scanning and treatment; CT scan is performed with the above position, the position is from the thoracic entrance to the horizontal position of the rib angle, the image data is saved and registered after the scanning is completed, and the patient's treatment position is determined, and the corresponding treatment is performed. The three-dimensional structure of the site was reconstructed and the gross tumor volume (GTV), clinical target volume (CTV), planned target volume (PTV) and surrounding endangered organ (AOR) were delineated by CT images according to the Varian Eclipse DX planning system to determine the target area and endangered The relationship of organs. Three-dimensional conformal radiation therapy: According to the size and location of the tumor target area and the relationship with important organs, design a reasonable irradiation field for the patient through the direction of the field, adjust the field parameters,

reduce the dose of the surrounding important organs and reach the target area. Excellent dose distribution.

From the isodose distribution curve, IMRT is tighter than the 3DCRT high-dose curve, and the conformability is better. Both treatment methods meet the requirements of more than 95% PTV receiving prescription dose, and the two treatment methods PTV dose comparison tips The difference between the mean dose and the maximum dose group was significant,  $p < 0.05$ , and the difference between the minimum dose groups was not significant,  $p > 0.05$ .

According to statistics, there are more than 200,000 new cases every year in Europe, and the related deaths account for about 20%. In the United States, there are about 150,000 people who die of lung cancer every year. In the last 10 years, the overall mortality rate of lung cancer has not changed significantly; in China, the incidence and mortality of lung cancer have increased significantly. According to the statistics of the Ministry of Health, lung cancer ranks first in the total mortality rate of malignant tumors. 83/100,000. Among male and female malignant tumor-related deaths, lung cancer caused deaths. Lung cancer is one of the malignant tumors that endanger human health. It is extremely malignant and prone to early metastasis. The curative effect is poor and the cure rate is low. Usually about 80% of patients are in the middle and late stage when they find lung cancer, thus losing the operation. In the best period, radiotherapy, as one of the important means of comprehensive treatment of tumors, plays an important role in the treatment of advanced lung cancer. Studies have shown that the incidence of lung cancer is on the rise, and its 5-year survival rate is only about 10%. Non-small cell lung cancer is the most common type, accounting for more than 80% of the total proportion of lung cancer, and in all patients with non-small cell lung cancer. Most of the patients are already locally advanced as soon as they are diagnosed. Radiation therapy is the most commonly used topical treatment for patients who cannot undergo surgery or refuse surgery, and is also one of the methods for comprehensive treatment of lung cancer. However, conventional radiotherapy methods are less effective, with a 2-year clinical survival rate of 33% to 72%, a 5-year survival rate of only 0 to 43%, and a 5-year survival rate of locally advanced cases as low as 5% to 10%. %between. However, the main cause of radiotherapy failure in locally advanced non-small cell lung cancer is distant metastasis and inability to control recurrence. Conventional two-dimensional radiotherapy can not effectively concentrate the dose on the target area of the tumor, and the dose received by the organ is high, and the total dose is difficult to increase.

**Dosimetric Study of IMRT and VMAT:** The breast position is fixed by the breast bracket, the patient is placed in the supine position, the bracket is adjusted to fully lift the affected arm, and the bracket parameters are recorded. Use the lead or lead to calibrate the affected breast range. Images were acquired using a spiral CT with a scan range of 5 cm from the thoracic entrance to 5 cm below the rib angle and a scan layer spacing of 5 mm. The CT images taken are transmitted to the treatment planning system via the network for use in delineating the target area and making plans.

The definition of the target area is based on the ES-TRO standard. CTV: Includes the entire affected side of the breast. Front boundary: 0.5cm under the skin; posterior: pectoralis major, sternum, intercostal muscle, no more than chest wall tissue; head lateral: visible or accessible upper edge of breast tissue; lateral boundary: visible or accessible The lower edge of the breast tissue; the medial boundary: the median midline; the lateral boundary: the breast is reflexed to the outside. PTV: The

inside and outside of the CTV are placed 0.5cm outside, and the head side and the side of the foot are placed 0.7 to 1.0cm, and the front boundary is not more than 0.5cm. Risk organs include: heart, affected lung, contralateral lung, spinal cord, and healthy breast.

CI, HI, and IHI are very important indicators for evaluating radiation therapy technology. The results of the study showed that the CI value of the VMAT program was significantly lower than that of the tangential field-based IMRT technology, and the HI value was also worse than the IMRT technique, especially for patients with left breast cancer and small breasts, so the patients did not VMAT illumination technology is recommended. In other studies, VMAT technology showed better CI values than IMRT, but the heart, ipsilateral lung, and healthy breast were exposed to higher doses. In addition, Park et al. showed that the VMAT technique had a better CI value than the IMRT technique, but there was no significant difference in the HI value, but the range of the low-dose area of the ipsilateral lung and heart was significantly increased. Our results and the above results reached a similar conclusion: there was no statistically significant difference in CI, HI and IHI between VMAT and IMRT. However, in patients with a small breast mass group, IMRT showed an advantage in CI values compared to VMAT.

The efficacy of conventional radiotherapy for locally advanced lung cancer is poor. The 5-year survival rate is 5% to 10%. It may be due to radiation injury in normal tissues such as radiation pneumonia and esophagitis, which significantly limits the higher dose on the target area of lung cancer. Shape radiotherapy technology has developed rapidly, and many units have developed this technology and used it for clinical treatment of lung cancer. Preliminary results show that the three-dimensional conformal radiotherapy of lung cancer has obvious dose distribution advantages, the target area is highly conformable, and the radiation damage of normal tissues is significantly reduced. This study compared three different radiotherapy plans in the same patient and found that the conformality of the conformal radiotherapy target area was significantly higher than that of conventional radiotherapy, while the volume percentage of total lung irradiation  $\geq 20$  Gy, the average dose of spinal cord, and esophageal irradiation.  $\geq 50$  Gy volume percentage, cardiac irradiation  $\geq 35$  Gy volume percentage, three-dimensional conformal radiotherapy was significantly lower than conventional radiotherapy, the difference was significant ( $P < 0.05$ ), confirmed that lung cancer three-dimensional conformal radiotherapy has obvious dose distribution advantages, It is possible to increase the dose of the lung cancer target area without increasing the radiation damage of the normal tissue to improve the therapeutic effect. In the design of a three-dimensional conformal radiotherapy program for lung cancer, how to define the target area is very important. It is difficult to make the target area accurate and not missed. In the plan design, it is found that the target area is enlarged by 1 cm, the normal tissue load is obviously increased, and the planning design difficulty is obviously increased. At present, the definition of lung cancer GTV is relatively clear, and it is a tumor seen on CT or MRI imaging, including lymph nodes larger than 1 cm in the mediastinum.

## CONCLUSION

Due to the development of society, people's requirements for quality of life have improved, and radiation therapy after breast-conserving surgery should reduce the treatment complications while ensuring local control rate. How to improve the uniformity of dose distribution in the target area and reduce the exposure dose of important organs such as heart, lung and contralateral breast have become the focus of discussion. IMRT and VMAT technology are effective tools for achieving better dose distribution and more effective protection of normal organs. Three-dimensional conformal radiotherapy is a new development in the current treatment of tumors. Based on the reconstruction of human body structure, the high-dose distribution is consistent with the shape of the target area under the guidance of the three-dimensional radiology plan. In addition, the intensity of the ray can be further adjusted so that rays of different intensities penetrate the target area to form an illumination field, separating the tumor from adjacent important tissues.

## References

- [1] Bentayeb, D., Lahrichi, N. and Rousseau, L.M., 2018. Patient scheduling based on a service-time prediction model: a data-driven study for a radiotherapy center. *Health care management science*, pp.1-15.
- [2] Mayo, C.S., Kessler, M.L., Eisbruch, A., Weyburne, G., Feng, M., Hayman, J.A., Jolly, S., El Naqa, I., Moran, J.M., Matuszak, M.M. and Anderson, C.J., 2016. The big data effort in radiation oncology: Data mining or data farming?. *Advances in radiation oncology*, 1(4), pp.260-271.
- [3] Zhang, S., Wang, H., Huang, W. and You, Z., 2018. Plant diseased leaf segmentation and recognition by fusion of superpixel, K-means and PHOG. *Optik*, 157, pp.866-872.
- [4] Mahmoodi, S.A., Mirzaie, K. and Mahmoudi, S.M., 2016. A new algorithm to extract hidden rules of gastric cancer data based on ontology. *SpringerPlus*, 5(1), p.312.
- [5] Ge, H., Yan, Y., Wu, D., Huang, Y. and Tian, F., 2018. Potential role of linc00996 in colorectal cancer: a study based on data mining and bioinformatics. *OncoTargets and therapy*, 11, p.4845.
- [6] Coroller, T.P., Agrawal, V., Huynh, E., Narayan, V., Lee, S.W., Mak, R.H. and Aerts, H.J., 2017. Radiomic-based pathological response prediction from primary tumors and lymph nodes in NSCLC. *Journal of Thoracic Oncology*, 12(3), pp.467-476.
- [7] Bakshi, S., Sa, P.K., Wang, H., Barpanda, S.S. and Majhi, B., 2018. Fast periocular authentication in handheld devices with reduced phase intensive local pattern. *Multimedia Tools and Applications*, 77(14), pp.17595-17623.
- [8] Mayo, C.S., Yao, J., Eisbruch, A., Balter, J.M., Litzenberg, D.W., Matuszak, M.M., Kessler, M.L., Weyburn, G., Anderson, C.J., Owen, D. and Jackson, W.C., 2017. Incorporating big data into treatment plan evaluation: Development of statistical DVH metrics and visualization dashboards. *Advances in radiation oncology*, 2(3), pp.503-514.