Proceedings of the 50th Zakopane School of Physics, Zakopane, Poland, May 18–23, 2015

Validation of the Usefulness of Dose Calculation Algorithms in Radiotherapy Planning System

B. Kiełtyka[a,*], K. Rawojć[a], K. Kisielewicz[b] and I. Markiewicz[c]
[a]Marian Smoluchowski Institute of Physics, Jagiellonian University, Cracow, Poland
[b]Maria Skłodowska-Curie Memorial Cancer Centre and Institute of Oncology Kraków Branch, Cracow, Poland
[c]Maria Skłodowska-Curie Memorial Cancer Centre and Institute of Oncology, Warsaw, Poland

One of the main goals of radiotherapy is to achieve tumor control and minimize probability of normal-tissue complications. For this reason radiation oncology requires high accuracy, which implies no more than 2–3% uncertainty levels in the treatment planning calculations. That is challenging, when heterogeneous tissues such as lungs and bones are involved. To verify the accuracy of the dose calculation algorithms numerous approaches might be performed. The most common are point dose, one-dimensional profile and two-dimensional isodose line comparison with experimental measurements. In presented study, results of transport modeling and the deposited spatial distribution of the dose, obtained by anisotropic analytical algorithm and pencil beam convolution algorithm, were compared to measurements recorded during the experiment. To achieve meaningful conclusions, three parameters: dose difference, distance to agreement and gamma parameter (γ) were taken into consideration and examined. The irradiation was performed using CIRS anthropomorphic phantom. For dose detection gafchromic EBT films were used and scanned after exposure using Epson Scanner. Measured and planned dose distributions were analyzed via FilmQA software. Preliminary results showed that the anisotropic analytical algorithm, with its complex accounting of heterogeneities, provides more accurate dose calculation within an area of a high density gradient, than pencil beam convolution does. The level of the data accuracy derived from the experiment was: dose difference (5%) — 83.4% and 68% pixels passing, distance to agreement (3 mm) — 99.0% and 96.7%, gamma parameter (for dose difference (3%), distance to agreement (3 mm)) — 90% and 75.5%, respectively, for anisotropic analytical algorithm and pencil beam convolution algorithms. The comparison between studied parameters dose difference, distance to agreement and γ for both algorithms implicated anisotropic analytical algorithm as an appropriate approach in radiotherapy treatment planning.

DOI: 10.12693/APhysPolA.129.219
PACS: 87.55.km, 87.55.kd, 87.10.Rt, 87.10.Vg

1. Introduction

One of the main goals of radiotherapy is to achieve tumor control and minimize probability of normal-tissue complications. This is the reason why radiation oncology requires high accuracy, which implies no more than 2–3% uncertainty levels in the treatment planning calculations [1]. That is challenging, when heterogeneous tissues such as lungs and bones are involved [2, 3]. To verify the accuracy of the dose calculation algorithms numerous approaches might be considered. Currently, numerous techniques are used to irradiate the tumors, as three-dimensional conventional radiation therapy (3D-RT), intensity-modulated radiation therapy (IMRT), tomotherapy, particle therapy and volumetric-modulated arc therapy (VMAT) [4]. Technology progress provides successive generations of treatment planning systems (TPS) for radiotherapy, which include new dose calculation algorithms and allow new irradiation techniques. These algorithms compute the dose for a given technique, subsequently showing the results as dosimetric parameters and displaying dose volume histograms or spatial isodoses [4]. In this article recently used algorithms like pencil beam convolution (PBC) and its modification, the analytical anisotropic algorithm (AAA), were taken into consideration and compared.

PBC algorithm assumes that the beam is divided into smaller elementary beams, which enter the irradiated volume. The dose distribution of whole beam is calculated by summing the dose distributions for each elementary beam [5] and the relative value of the energy fluence in air is uniform across irradiated volume (omitting the boundary effects). In majority of clinical applications, calculations of dose to the tumor are performed using PBC algorithm because of its simplicity and rapidity [6]. Nevertheless, PBC algorithm has shortcomings when it comes to severe inhomogeneities, particularly in regions where charged particle equilibrium does not exist. Therefore it is questionable for target dose calculations for i.e. lung cancer treatments.

The introduction of convolution-superposition (CS) algorithms that better account for electron transport, have enabled improved calculation of dose distribution, principally in the absence of electronic equilibrium [6]. In the Eclipse TPS (Varian Medical Systems) the AAA is implemented. The AAA is a further 3D development of PBC-based superposition algorithm [7]. It uses Monte Carlo simulations of separately modeled contributions from three sources: primary photons, extra-focal photons and contaminating electrons; each of these has an

*corresponding author; e-mail: bartosz.kieltyka@uj.edu.pl

associated fluence, an energy deposition density function and a scatter kernel. Monte Carlo (MC) simulation is considered to be a gold standard in dose calculation, and it is therefore used to evaluate other dose calculation algorithms [6, 7]. Lately relevant adjustments associated with modeling of tissue heterogeneity, increasing the accuracy of dose calculation of the scattered radiation or new generations of medical equipment have been made to AAA in order to achieve higher quality of dose deposition calculations.

This study presents a series of statistical tests implemented in a step by step procedure that may be used by medical physicist in order to compare the dosimetric outcome of two different dose calculation algorithms. The results of transport modeling and the deposited spatial distribution of the dose, obtained by using different algorithms, were compared to measurements recorded during the experiment. To achieve meaningful conclusions, three parameters: dose difference (DD), distance to agreement (DTA) and gamma parameter (γ) were taken into consideration and examined. The procedure is presented with an emphasis on the application in radiotherapy rather than the underlying mathematical principles, which are not detailed.

2. Materials and methods

In our study, results of transport modeling and the deposited spatial distribution of the dose, obtained by AAA and PBC algorithm, were compared to measurements recorded during the experiment. Treatment plans were prepared in the Eclipse 8.1 TPS (Varian Medical Systems). The dose distributions of the clinical treatment plans initially performed using the PBC algorithm were recalculated with AAA using the same plan parameters provided by PBC.

The treatment was simulated using CIRS ATOM®701-706 anthropomorphic phantom which is commonly used for the dosimetric procedures. The photon irradiation was performed using conformal technique and only in the chosen region of high heterogeneity — human chest. For dose detection gafchromic EBT films (Ashland Inc.) were used and scanned after exposure via Epson Expression Scanner 10000XL with resolution 2400 dpi. The sensitivity of gafchromic EBT films was 1–800 cGy and the uncertainty level reaches no more than 2%. Measured and planned dose distributions were analyzed via FilmQA software.

3. Statistical analysis

The derived data quality was analyzed at hand to fulfill the assumption of banking parametric data. For those data sets which were distributed normally and had a similar variance between groups the paired Student t-test was used for statistical comparison of the means obtained by using two algorithms. If the data do not fulfill main assumption, alternative nonparametric tests were used such as the Wilcoxon–Mann–Whitney and the Fisher exact test. The non-parametric Wilcoxon rank test takes into account the signed-rank of the difference between each pair of measures instead of using all the absolute data [4]. The test does not require a normal distribution and does not considers the size of the difference. These features make the Wilcoxon signed-rank test a good match for radiotherapy data analysis, since these data are normally paired according to the possibility to generate multiple different results for each medical case. All tests were two-tailed with a p value of < 0.05 considered the threshold for statistical significance. To achieve meaningful conclusions, three parameters: dose difference (DD), distance to agreement (DTA) and gamma parameter (γ) were taken into consideration and examined using StatSoft®program.

4. Results

To perform the Fisher exact test all data sets were divided into three subgroups: A — 95%, B — 90%, C — 80%. The groups were tested for pixels compliance from PBC and AAA. Obtained values are presented as a percentage of pixels compliance showing a correlated match between treatment plan (modeled by the dose calculation algorithms) and actual dose distribution for separate parameters measured during the experiment. The accepted threshold is defined by each parameter accordingly.

The first step of analysis showed a clear shift in the mean value of pixels compliance calculated by AAA compared to values obtained by PBC (Table I).

TABLE I

The results of the mean pixel value, standard deviation and median presented as a percentage of pixels fulfilling the assumed criteria of each tested parameter, respectively. The results are comparison of values obtained by using AAA and PBC to those recorded during the experiment.

Tested parameter	Results obtained by using PBC			Results obtained by using AAA			Statistical test
	Mean	SD	Median	Mean	SD	Median	
Gamma (DD = 3%, DTA = 3 mm)	75.5	7.8	78.2	90.0	8.3	93.6	U-test
Gamma (DD = 4%, DTA = 3 mm)	76.7	10.7	–	87.8	9.6	–	T-test
DD (5%)	68.0	16.3	71.1	83.4	15.0	89.7	U-test
DD (3%)	48.1	12.5	–	61.8	17.5	–	T-test
DTA (3 mm)	96.7	2.7	97.2	99.0	1.3	99.4	U-test
DTA (2 mm)	88.9	5.0	–	96.2	3.2	–	T-test

The distance of the mean value is 14.5% for γ parameter (DD = 3%, DTA = 3 mm). Increase of the number of observations did not show statistically significant difference in the tested hypothesis.

The results further confirm that using AAA algorithm we are able to estimate the actual dose distribution with higher precision which is shown in Table II.

The results of the Fischer exact test presented a characteristic increased trend in the number of data matching the criteria of binary division. In case of AAA even a

TABLE II

The statistical comparison of the results obtained by using PBC and AAA for the gamma parameter, DD, DTA.

The results of T-test for independent variables					
Tested parameter	Mean PBC	Mean AAA	t^a	df^b	p^c
Gamma (DD = 4%, DTA = 3 mm)	76.7	87.8	–2.56	20	0.018
DD (3%)	48.1	61.8	–2.12	20	0.047
DTA (2 mm)	88.9	96.2	–4.07	20	0.00059
U-test of Mann–Whitney					
Tested parameter	The sum of ranks PBC	The sum of ranks AAA	U^d	Z^e	p^c
Gamma (DD = 3%, DTA = 3 mm)	80	173	14	–3.02	0.0025
DD (5%)	90	163	24	–2.36	0.018
DTA (3 mm)	82	171	16	–2.89	0.0039

[a]the results of T-test, [b]number of degrees of freedom, [c]significance level — "p-value", [d]the results of U-test, [e]the results of Z-test (for the Mann–Whitney U-test).

TABLE III

The statistical comparison of the results obtained by using the Fischer exact test for two gamma parameter variants.

The results of the Fischer exact test		
Subgroup	Gamma (DD = 3%, DTA = 3 mm)	Gamma (DD = 4%, DTA = 3 mm)
95%	$p = 0.0175$	$p = 0.1071$
90%	$p = 0.0062$	$p = 0.0062$
80%	$p = 0.0150$	$p = 0.3297$

slight change in criteria grouping is noticeable. PBC begins to fulfill the given criteria only by lowering the threshold to 80% in the pixels line. The p value calculated for the γ parameter (DD = 3%, DTA = 3 mm) in all analyzed criteria clearly indicates the existence of a difference between the treatment algorithms. The AAA algorithm provides more adequate dose distribution (p-values significantly below 0.05) than PBC does. The extension of accepting criteria to small extent (about 1%) shows that the results are starting to be similar for both algorithms (see Table II and Table III). As shown in Table III only at 90% criterion the Fischer exact test pointed out the vast differences of compared dose calculation algorithms in favor of AAA.

5. Conclusions

Among the various statistical approaches available, the medical physicist has to make a choice adapted to the particularities of radiation therapy of each medical case. Due to this, the particular nature of the data and the way they derived the quality of each set needs a deep analysis. One should assess whether the data at hand fulfill the assumptions of the parametric data, i.e., are distributed normally or have similar variance between groups. When the data fulfill these assumptions, Student t-test or one way ANOVA might be used to compare the means, whereas non-parametric Wilcoxon signed-rank test should be used if the dose differences do not fulfill the assumptions of parametric data.

PBC substantially overestimates the dose to the tumor, while the AAA is more similar to the MC simulation. Therefore it is recommended that the treatment plans for regions of higher heterogeneity should be developed using an advanced dose calculation algorithm. The PBC algorithm might give satisfactory results when it comes to providing treatment in homogeneous regions of the human body i.e. brain. Each algorithm has its advantages in certain applications.

This paper illustrates and justifies the use of statistical tests for dosimetric comparisons in radiotherapy. The statistical analysis shows the significance of dose differences resulting from two techniques of dose calculations in radiotherapy. The comparison of the two algorithms in the present study is in accordance with the literature.

References

[1] I.M. Gagné, S. Zavgorodni, *J. Appl. Clin. Med. Phys.* **8**, 33 (2006).

[2] International Commission on Radiation Units and Measurements (ICRU), *Determination of absorbed dose in a patient irradiated by beams of X or gamma rays in radiotherapy procedures*, ICRU Report 24, Washington (DC) 1976.

[3] A. Van Escha, E. Namur, L. Tillikainen, J. Pyykkonen, M. Tenhunen, H. Helminen, S. Siljamäki, J. Alakuijala, M. Paiusco, M. Iori, D.P. Huyskens, *Med. Phys.* **33**, 4130 (2006).

[4] A. Chaikh, J.Y. Giraud, E. Perrin, J.P. Bresciani, J. Balosso, *Radiat. Oncol.* **9**, 205 (2014).

[5] A. Ahnesjö, M. Saxner, A. Trepp, *Med. Phys.* **19**, 263 (1992).

[6] Y. Zhao, G. Qi, G. Yin, X. Wang, P. Wang, J. Li, M. Xiao, J. Li, S. Kang, X. Liao, *Radiat. Oncol.* **9**, 287 (2014).

[7] W. Ulmer, J. Pyyry, W. Kaissl, *Phys. Med. Biol.* **50**, 1767 (2005).