

## Original Research Article

## Comparison of complexity metrics for multi-institutional evaluations of treatment plans in radiotherapy

Victor Hernandez<sup>a,1,\*</sup>, Jordi Saez<sup>b,1</sup>, Marlies Pasler<sup>c</sup>, Diego Jurado-Bruggeman<sup>d</sup>, Nuria Jornet<sup>e</sup><sup>a</sup> Department of Medical Physics, Hospital Universitari Sant Joan de Reus, IISPV, Tarragona, Spain<sup>b</sup> Department of Radiation Oncology, Hospital Clínic de Barcelona, Spain<sup>c</sup> Lake Constance Radiation Oncology Center Singen-Friedrichshafen, Germany<sup>d</sup> Medical Physics and Radiation Protection Department, Institut Català d'Oncologia, Girona, Spain<sup>e</sup> Servei de Radiofísica i Radioprotecció, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

## ARTICLE INFO

## Keywords:

Plan complexity  
Complexity metrics  
Audits  
Clinical trials

## ABSTRACT

**Background and purpose:** It is known that intensity-modulated radiotherapy plans that are highly complex might be less accurate in dose calculation and treatment delivery. Multiple complexity metrics have been proposed, but the relationships between them have not been thoroughly investigated. This study investigated these relationships in multi-institutional comparisons of treatment plans, where plans from multiple treatment planning systems (TPSs) are typically evaluated.

**Materials and methods:** A program was developed to compute several complexity indices and provide analysis of dynamic plan parameters. This in-house software was used to analyse plans from a recent multi-institutional audit. Additionally, 100 clinical volumetric modulated arc therapy (VMAT) plans from two institutions using different TPSs were analysed.

**Results:** All plans produced satisfactory pre-treatment verification results and, hence, complexity metrics could not be used to predict plans failing QA. Regarding the relationship among complexity indices, some very strong correlations were found ( $r > 0.9$  with  $p < 0.01$ ). However, some relevant discrepancies between complexity indices were obtained, even with negative correlation coefficients ( $r \sim -0.6$ ) which were expected to be positive. These discrepancies could be explained because each complexity index focused on different features of the plan and different TPSs prioritised modulation of different plan parameters.

**Conclusions:** Some complexity indices provided similar information and can be considered equivalent. However, indices that focused on different plan parameters yielded different results and it was unclear which complexity index should be used. Careful consideration should be given to the use of complexity metrics in multi-institutional studies.

## 1. Introduction

Advances in the technology for planning and delivery of radiotherapy treatments allow for highly conformal dose distributions to be achieved. However, these distributions require modulation of many machine parameters [1–5]. Since additional sources of variability are thus introduced, treatment plans with similar dose distributions may differ greatly in their complexity. Many investigators have reported that the degree of plan complexity may affect the accuracy of dose calculations and treatment delivery [6–13], which is crucial in dosimetry audits and clinical trials, as well as for big data analysis [14–16]. Therefore, aspects such as quality and complexity of treatment plans have to be carefully evaluated in multi-institutional plan comparisons

[17].

Several investigators have proposed different complexity metrics and have reported correlations with overall accuracy and the resulting quality assurance (QA) metrics [6–13]. Thus, less complex plans offer several benefits such as more accurate dose calculations, more accurate and robust treatment delivery, better QA metrics and even lower risk of intra-fraction movements and patient variations [6–10]. For all these reasons plans with low complexity are associated with lower uncertainties and can be considered, in general, more robust than highly complex plans.

AAPM pointed out the need to incorporate measures of beam modulation to ensure that centres achieve intensity-modulated radiation-therapy (IMRT) plans that are comparable with regards to their

\* Corresponding author.

E-mail address: [vhernandezmasgrau@gmail.com](mailto:vhernandezmasgrau@gmail.com) (V. Hernandez).<sup>1</sup> Both authors equally contributed to this study.

complexity [18]. However, it is not clear which of the proposed complexity indices should be used and the relationship between these multiple indices in multi-TPS environments has not been previously addressed. In this study we investigated the use of complexity metrics in multi-institutional comparisons where multiple TPSs, planners and linac types are typically involved. The study focused on volumetric modulated arc therapy (VMAT) treatments, but most of the indices evaluated can also be applied to other techniques such as sliding window and step-and-shoot IMRT.

## 2. Materials and methods

### 2.1. Complexity metrics

In this study several complexity indices that are computed from the treatment plan parameters defined at each of the control points of the plan were investigated. These indices allow for a detailed analysis of the dynamic parameters involved in treatment plans, which makes them more appropriate for VMAT than ‘fluence-based’ indices. The following indices were evaluated:

- a) Modulation Complexity Score (MCS) [6]. This score integrates two contributions to complexity: variability in the shape of segments and variations in their area. MCS uses a fixed range from 0 to 1 and, unlike the rest of the complexity indices, it is defined in such a way that the lower the value of the MCS the higher the complexity. It was initially designed for step-and-shoot treatments and later adapted to sliding window and VMAT [9,19].
- b) Edge metric (EM) [8]. This metric computes the complexity of multileaf collimator (MLC) apertures based on the ratio of MLC side edge length and aperture area. In this study the original recommendation for the parameters ( $C1 = 0$  and  $C2 = 1$ ) was followed. Thus, the greater the differences between the positions of adjacent leaves the higher the EM index, which is closely related to the amount of tongue-and-groove effect.
- c) Leaf travel (LT) [9]. This index indicates the average distance travelled by the moving leaves. LT was devised for VMAT treatments consisting of a single full arc. To allow for simple comparisons between plans with a different number of arcs or with partial arcs, we divided LT by the corresponding arc length (typically about 360 deg for single arcs and about 720 deg for double arcs) and we named this index as ‘LT/AL’.
- d) Plan irregularity (PI) and Plan modulation (PM) [11]. PI describes the deviations of aperture shapes from a circle, being 1 for a perfect circle. PM indicates to what extent a beam is modulated with multiple smaller segments.
- e) Modulation index total (MI<sub>total</sub>) [10]. This index evaluates the variations in speed and acceleration of the MLC as well as variations of the gantry speed and the dose rate. MI<sub>total</sub> is, to our knowledge, the only complexity index that takes into account the modulation of the dose rate and the gantry speed.

### 2.2. Treatment plans evaluated

The first group of plans evaluated in this study included forty plans from a recent audit promoted by the Catalan Association of Medical Physics within the framework of the Catalan-Occitan Oncology Group (GOCO). This audit included local pre-treatment verification results and independent dosimetry audit measurements [20]. A mock head-and-neck and a mock prostate case adapted from those proposed in TG119 were used. Most plans (twenty-eight) were produced with Eclipse™ (Varian Medical Systems), eight plans were generated with Pinnacle Auto-Planning (Philips Radiation Oncology Systems) and four plans with Monaco (Elekta AB). Hereafter these TPSs will be called TPS-A, TPS-B and TPS-C, respectively. Details on the TPSs, the linacs and the methodology used can be found in the aforementioned publication.

Additionally, in the present study clinical plans from TPS-A and TPS-B were also analysed. In particular, fifty head-and-neck VMAT plans and fifty prostate VMAT plans from each TPS were randomly selected and evaluated. Plans from TPS-A and TPS-B were produced for a Varian Clinac iX (Millennium 120 MLC) and an Elekta Synergy (MLCi2, binned dose rate), respectively.

### 2.3. Software and equipment used

To compute the previously described complexity indices, an in-house program called PlanAnalyser was developed in MATLAB (Mathworks, Massachusetts, USA). This software reads the DICOM plan as exported from the TPS and computes complexity indices using the data contained in the DICOM plan. Plan complexity indices were computed by joining all beams and performing the calculations for the ‘combined’ beam.

PlanAnalyser incorporates an emulator that predicts the variations of the dynamic plan parameters during treatment delivery. Since one of the complexity metrics (MI<sub>total</sub>) evaluates the variations of the dose rate and gantry speed, we investigated the modulation of these parameters. Mean variations were defined as the total variation (i.e., sum of all variations between consecutive control points) divided by the total arc length. To verify the predictions from the emulator they were compared to results from log files for both Elekta and Varian linacs. Varian log files were analysed with in-house software [21] and log files from Elekta were recorded with the service graphing module of the linac controller (Integrity 1.2).

Pre-treatment verifications were carried out for all plans. Audit plans were measured with both independent QA equipment (ArcCHECK, Sun Nuclear Corporation) and a large variety of local QA devices [20]. Clinical plans from TPS-A and TPS-B were measured with ArcCHECK and Octavius II – 2D array seven29 (PTW Freiburg), respectively. Since audit plans corresponded to the same mock cases, a plan quality score was computed with the software PlanIQ™ (Sun Nuclear Corporation) in order to identify which plans achieved the best trade-off between target coverage, homogeneity, conformity, and doses to organs at risk [20].

To investigate the dependencies among these indices, the Spearman’s rank correlation coefficients  $r$ , sensitive to both linear and non-linear correlations, were calculated. The strength of the association, for absolute values of  $r$ , 0–0.19 was regarded as ‘no correlation’, 0.20–0.39 as ‘weak’, 0.40–0.59 as ‘moderate’, 0.60–0.79 as ‘strong’ and 0.80–1 as ‘very strong’. To account for multiple testing, false discovery rates (q-values) [22,23] were calculated. Reported p-values represent statistical analysis without multiple testing correction and statistical significance was considered at  $p < 0.05$  with q-value  $< 0.1$ . All statistical analysis was performed in R-3.3.2 (R: A Language and Environment for Statistical Computing, 2016, Vienna, Austria).

## 3. Results

### 3.1. Audit plans

All participating centres fulfilled all the requested planning goals regarding both target coverage requirements and dose limits to organs at risk. Large differences in the degrees of plan complexity were observed, but no statistically significant correlation was found between dosimetric plan complexity and plan quality [20]. Pre-treatment verification results were clinically acceptable for all plans ( $> 95\%$  of points with  $\gamma\ 3\%/3\ \text{mm} < 1$ ), hence complexity metrics could not be used to predict plans failing QA.

Regarding the comparison between complexity indices, strong correlations were found between MCS, PI and EM. However, we also observed some evident discrepancies, meaning that some plans were more complex than others according to a particular complexity index, while the opposite result was found when another complexity index was

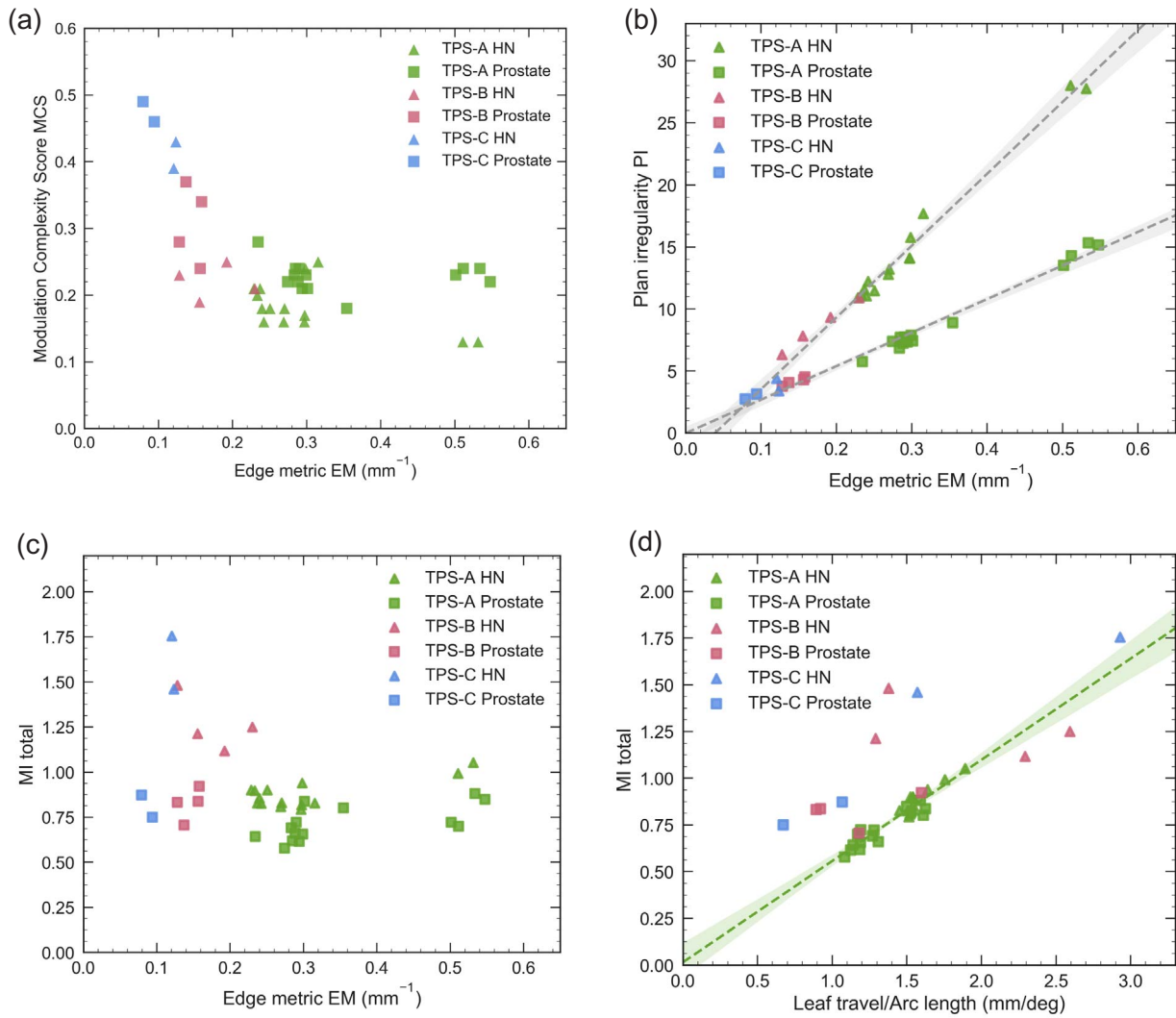


Fig. 1. Scatter plots for four pairs of complexity indices. Data corresponds to audit plans for the same mock cases. The shaded areas in (b) correspond to linear fits of the head and neck and prostate cases separately taking into account data from all treatment planning systems. The shaded area in (d) corresponds to a linear fit to the data for TPS-A plans regardless of the treatment site.

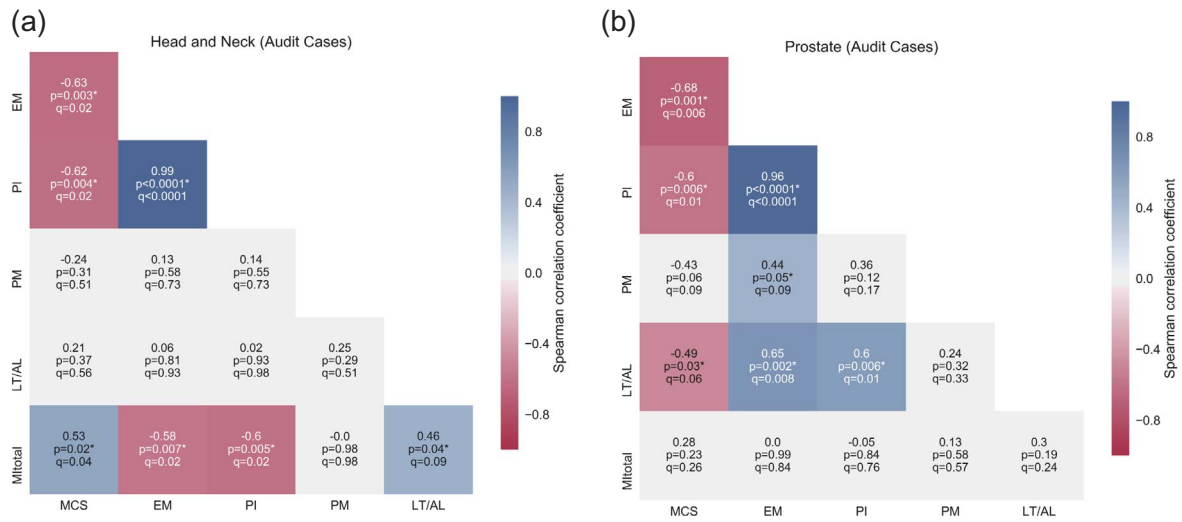


Fig. 2. Spearman's correlation coefficients and their statistical significance  $p$  and false discovery rates  $q$  for the (a) head-and-neck and (b) prostate plans from the audit. Statistically significant correlations are marked (\*) and coded in colour.

considered. In Fig. 1 the comparison between four pairs of indices is shown and a more detailed comparison between all pairs of indices is provided as [Supplementary material \(Fig. S1\)](#). Spearman's correlation coefficients are given in Fig. 2.

Fig. 1a shows an inverse correlation between MCS and EM. This was expected because lower MCS values are linked to higher plan complexities and results in negative correlation coefficients. Indeed, a strong correlation was found between MCS and EM, with  $r = -0.63$  ( $p = 0.003$ ) for head-and-neck and  $r = -0.68$  ( $p = 0.001$ ) for prostate mock cases. However, six plans from TPS-A yielded a high complexity according to EM (with values 0.5–0.55) while their complexity according to MCS was very similar to the other plans with the same TPS. All plans from TPS-A were more complex than plans from TPS-C according to these indices (i.e., higher EM and lower MCS values), while plans from TPS-B fell between them.

As shown in Fig. 1b, a very strong linear correlation was also found between PI (related to MLC aperture irregularity) and EM (related to the tongue-and-groove effect) for a given treatment site, with  $r > 0.95$  ( $p < 0.0001$ ). The ratio between the two indices depended on the treatment site, probably because the size of the target volumes was quite different and EM is not dimensionless. Thus, PI revealed higher complexities for head-and-neck plans than for prostate plans, while EM indicated similar complexities in both cases. According to PI, plans from TPS-C were also the least complex and plans from TPS-A were the most complex.

On the contrary, no statistically significant correlation between MI<sub>total</sub> and EM was found for prostate cases and a negative moderate correlation ( $r = -0.58$ ,  $p = 0.007$ ) was found for head-and-neck plans. In general, Fig. 1c shows that plans with higher EM values did not produce higher values of MI<sub>total</sub>. Interestingly, MI<sub>total</sub> provided similar or higher plan complexities for both TPS-B and TPS-C with respect to TPS-A, while EM yielded opposite results. As it can be observed, there was no evident correlation between MI<sub>total</sub> and EM for any TPS and the negative correlation coefficient for head-and-neck plans was caused by these differences between TPSs.

The relationship between MI<sub>total</sub> and LT/AL is illustrated in Fig. 1d. Although these two indices focus on different features of the plan, for plans from TPS-A a very strong correlation was obtained regardless of the treatment site, as the regression line in Fig. 1d shows ( $r > 0.99$ ,  $p < 0.0001$ ). On the contrary, when all TPSs were considered the overall correlation was much weaker and only statistically significant for the head-and-neck plans (see Fig. 2).

The variations in dose rate and gantry speed for plans from different TPSs are shown in Fig. 3. Variations for TPS-B and TPS-C were much larger than those produced by TPS-A, which indicates that the

optimisation engines in TPS-B and TPS-C further modulate both the dose rate and the gantry speed.

To confirm these variations and validate our predictions, log files were collected and analysed. Data from log files was in good agreement with predictions from the emulator and confirmed the different degree of modulation of these parameters depending on the TPS (see Fig. 4).

### 3.2. Clinical plans

Pre-treatment verification was carried out for all the clinical plans evaluated. All QA results were also clinically acceptable ( $> 95\%$  of points with  $\gamma 3\%/3\text{ mm} < 1$ ). Complexity indices from clinical plans were, in general, similar to those found in the audit, which confirmed that plans from the audit were representative of clinical practice. For clinical plans the variability in complexity scores was higher because they included a wide variety of cases with large anatomical variations, especially for head-and-neck cases.

The relationships between complexity indices and their correlations were similar to those found in the audit. A detailed comparison of all the indices and the correlations obtained is provided as [Supplementary material \(Figs. S2 and S3\)](#). For comparison purposes, the relationship between MI<sub>total</sub> and EM for clinical plans is shown in Fig. 5a and b. Discrepancies were very similar to those observed in the audit, with plans from different TPSs being more complex according to EM and less complex according to MI<sub>total</sub>. Again, plans from TPS-A involved much lower variations of dose rate and gantry speed, with practically no variations in gantry speed when multiple arcs were used (see Fig. 5c and d). TPS-B, on the contrary, produced much larger variations in both dose rate and gantry speed regardless of the number of arcs involved. Boxplots illustrating these differences are provided as [Supplementary material \(Fig. S4\)](#).

## 4. Discussion

We found very strong correlations between some of the complexity indices that evaluate similar parameters (e.g., MCS, PI and EM), which means they provide similar information. In general, indices that focused on different features of the plan produced much weaker correlations or no statistically significant correlations. Some discrepancies appeared between complexity metrics such as MI<sub>total</sub> and EM for both audit plans and clinical plans when plans produced by different TPSs were compared. These discrepancies can be explained by the large differences in the degree of modulation of the dose rate and the gantry speed depending on the TPS. Indeed, MI<sub>total</sub> is further increased in plans that further modulate the dose rate and the gantry speed, while the other

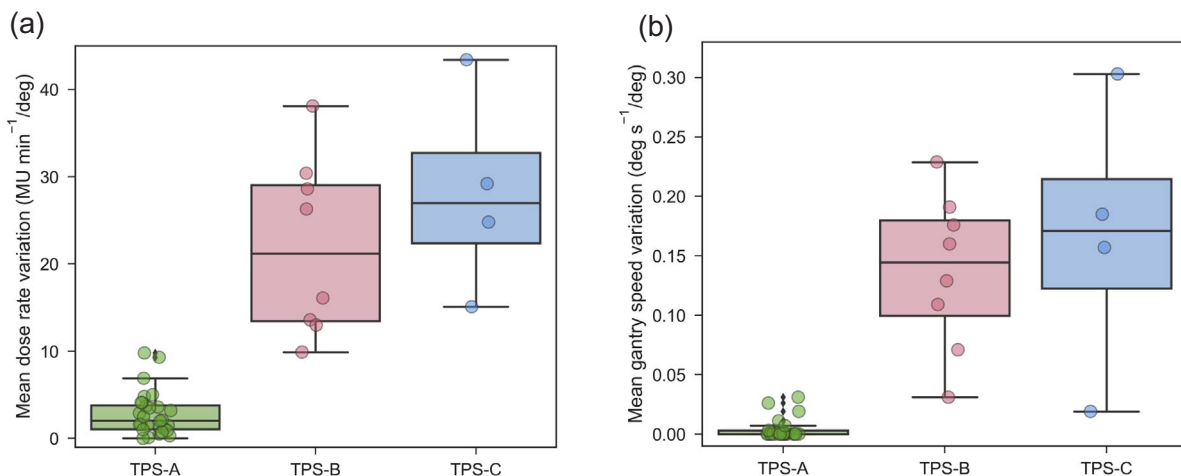


Fig. 3. Box plots showing (a) the mean dose rate variations and (b) the mean gantry speed variations for the audit plans and the three treatment planning systems evaluated. The central line indicates the median value, the box limits represent the 1st and 3rd quartile and the whiskers indicate the minimum and maximum values.

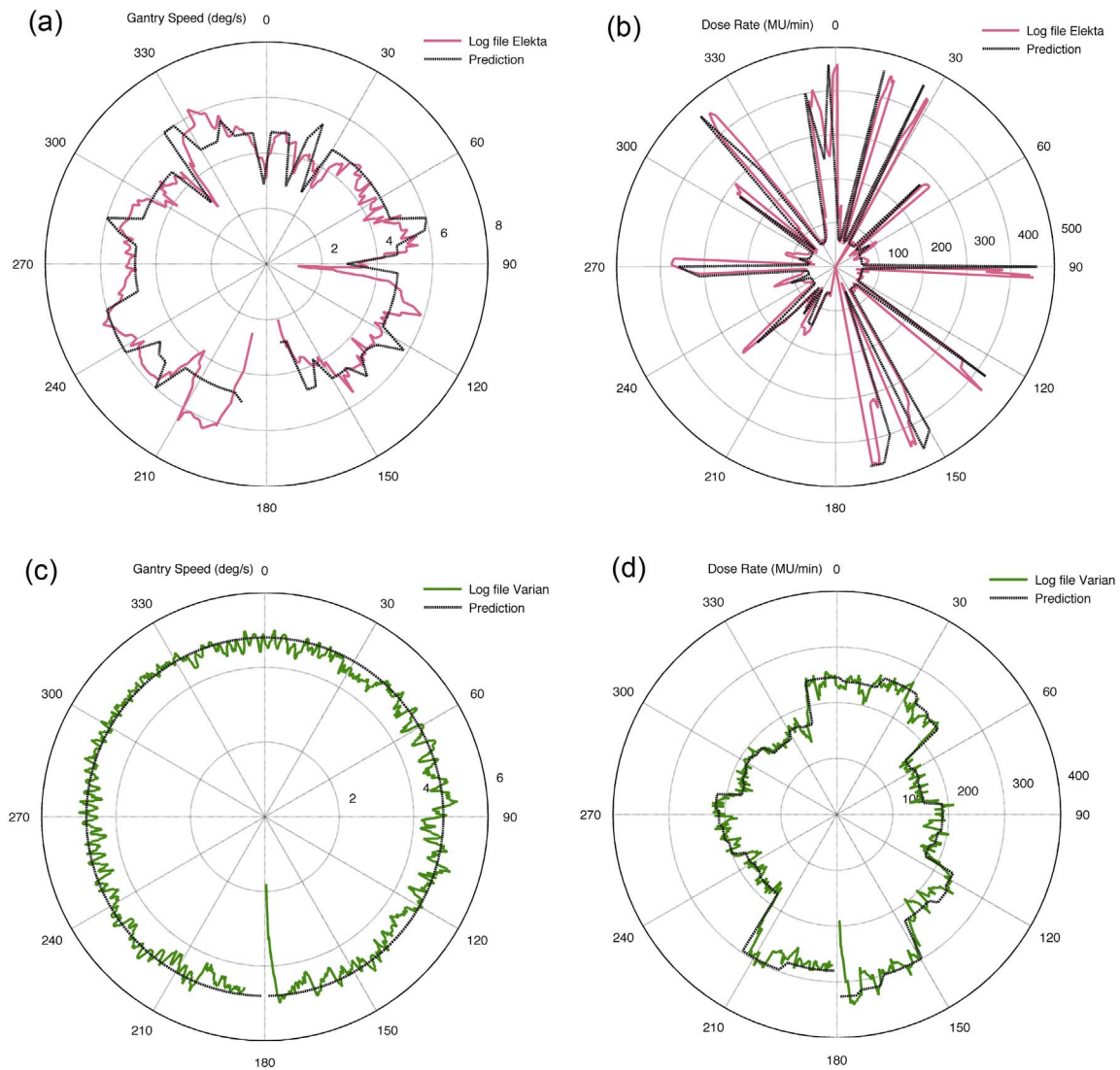


Fig. 4. Dose rate and gantry speed as a function of the gantry angle for plans from TPS-B (top) and TPS-A (bottom). Data from log files and predictions from the emulator are given for two representative plans corresponding to the same head-and-neck case.

complexity indices do not take these variations into account. The variability of complexity metrics depended on the TPS, which indicates that the ability of each index to discriminate between plans depends on the TPS. However, it is not clear which index is more relevant for each TPS or linac. This should probably be investigated in every particular situation and based on specific QA results.

The discrepancies between different complexity metrics indicate that no individual metric is sufficient for all TPSs and that several complexity metrics should be evaluated. Additionally, in order to identify excessively complex plans, multi-institutional comparisons should be carried out per TPS. Similarly, acceptable ranges or threshold levels in complexity metrics depend on the TPS model and hence must be evaluated for each specific TPS.

The analysis of complexity indices revealed differences in the optimisation engines and sequencers of each TPS, which prioritise the modulation of different plan parameters. A large spread in some of the complexity indices was observed depending on the TPS, in EM for TPS-A, for instance, and in MItotal for TPS-B and TPS-C, although all these plans achieved similar dosimetric plan quality. This indicates that more complex plans do not necessarily produce better dose distributions, which has been reported by other investigators [8,24]. A certain degree of plan complexity is not per se negative, because some degree of complexity is needed to achieve clinically acceptable dose distributions

[25]. However, unnecessary complexity should be avoided and excessively complex plans might compromise the accuracy of dose calculations in the TPS and the accuracy of treatment delivery [6–13]. For that reason some authors recommend incorporating complexity metrics into the cost function used by optimisation algorithms [6–8,24,26]. In our audit, we found several plans with a much higher degree of complexity that produced dose distributions of similar dosimetric quality. By incorporating these metrics into optimisation algorithms, the degree of modulation of dynamic plan parameters could be further controlled and this unnecessary complexity might be greatly reduced.

The fact that plans produced by different TPSs prioritise the modulation of different plan parameters can also have implications in the commissioning and QA of linacs and TPSs. For instance, plans from TPS-A were more demanding for the MLC, which makes them more sensitive to uncertainties associated to the use of small MLC openings, as well as to potential errors in the MLC calibration. On the other hand, plans from TPS-B and TPS-C further modulated the gantry speed and the dose rate and were, therefore, more demanding for the gantry assembly and beam modulation. Complexity metrics can be useful to understand these differences and to better adapt QA programs to each particular situation.

Many investigators have reported that plan complexity indices are correlated with QA metrics [6–13]. However, in this study all plans

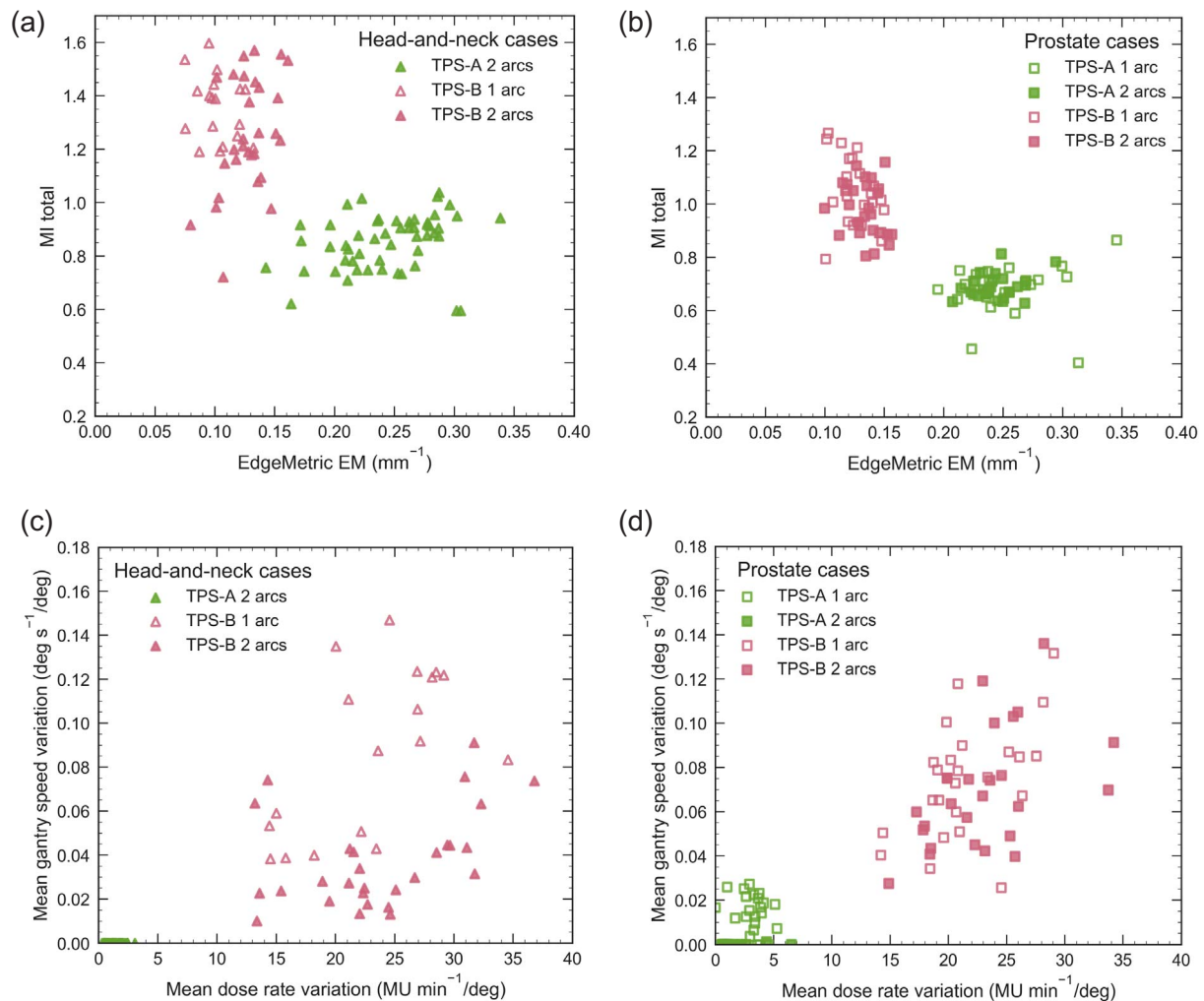


Fig. 5. The comparison between MI total and EM is shown for (a) head-and-neck and (b) prostate clinical plans. In (c) and (d) the corresponding mean gantry speed and mean dose rate variations are shown.

produced satisfactory QA results and more stringent QA criteria might be needed to investigate these potential correlations. If multi-institutional comparisons (such as in audits and clinical trials) include some plans that fail QA, complexity metrics can indicate if these plans are particularly complex, which might help understand the causes for poor QA results. In general, these causes may depend on the linac model and its proper maintenance, as well as on the limitations of the TPS used (beam model, dose engine and potential commissioning inaccuracies). Since different causes would originate different correlations between QA results and complexity metrics, these correlations are not generic and might be harder to find in multi-institutional comparisons, where different causes for poor QA results could interfere.

One limitation of this study is that plans from only a few TPSs were evaluated. We found, however, clear differences in the degree of modulation of their plan parameters and in some of their corresponding complexity metrics. Another limitation is that only VMAT plans were analysed, but the software developed can also be applied to other IMRT techniques.

In conclusion, strong correlations were found between several complexity metrics, which show that some indices provide similar information and can be considered equivalent. However, some relevant discrepancies between complexity metrics were also found and it is unclear which complexity index should be used. The ranking of plans according to their degree of complexity greatly depends on the metric used and on the features evaluated by each index, especially for plans

from different TPSs. This must be carefully considered in multi-institutional plan comparisons, such as audits and clinical trials.

#### Conflict of interest statement

None declared.

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.phro.2018.02.002>.

#### References

- [1] Ling CC, Zhang P, Archambault Y, Bocanek J, Tang G, LoSasso T. Commissioning and quality assurance of RapidArc radiotherapy delivery system. *Int J Radiat Oncol Biol Phys* 2008;72(2):575–81.
- [2] Van Esch A, Huyskens DP, Behrens CF, Samsøe E, Sjolín M, Bjelkengren U, et al. Implementing RapidArc into clinical routine: a comprehensive program from machine QA to TPS validation and patient QA. *Med Phys* 2011;38(9):5146–66.
- [3] Thwaites D. Accuracy required and achievable in RT dosimetry: have modern technology and techniques changed our views? *J Phys Conf Ser* 2013;444:012006.
- [4] van der Merwe D, Van Dyk J, Healy B, Zubizarreta E, Izewska J, Mijnheer B, et al. Accuracy requirements and uncertainties in radiotherapy: a report of the International Atomic Energy Agency. *Acta Oncol* 2017;56(1):1–6.
- [5] Li JS, Lin T, Chen L, Price Jr RA, Ma CM. Uncertainties in IMRT dosimetry. *Med Phys* 2010;37:2491–500.
- [6] McNiven AL, Sharpe MB, Purdie TG. A new metric for assessing IMRT modulation complexity and plan deliverability. *Med Phys* 2010;37(2):505–15.

- [7] Nauta M, Villarreal-Barajas JE, Tambasco M. Fractal analysis for assessing the level of modulation of IMRT fields. *Med Phys* 2011;38(10):5385–93.
- [8] Younge KC, Matuszak MM, Moran JM, McShan DL, Fraass BA, Roberts DA. Penalization of aperture complexity in inversely planned volumetric modulated arc therapy. *Med Phys* 2012;39(11):7160–70.
- [9] Masi L, Doro R, Favuzza V, Cipressi S, Livi L. Impact of plan parameters on the dosimetric accuracy of volumetric modulated arc therapy. *Med Phys* 2013;40(7):071718.
- [10] Park JM, Park S, Kim H, Kim JH, Carlson J, Ye SJ. Modulation indices for volumetric modulated arc therapy. *Phys Med Biol* 2014;59(23):7315–40.
- [11] Du W, Cho SH, Zhang X, Hoffman KE, Kudchadker RJ. Quantification of beam complexity in intensity-modulated radiation therapy treatment plans. *Med Phys* 2014;41(2):21716.
- [12] Götstedt J, Hauer AK, Bäck A. Development and evaluation of aperture-based complexity metrics using film and EPID measurements of static MLC openings. *Med Phys* 2015;42:3911–21.
- [13] Crowe SB, Kairn T, Kenny J, Knight RT, Hill B, Langton CM, et al. Treatment plan complexity metrics for predicting IMRT pre-treatment quality assurance results. *Australas Phys Eng Sci Med* 2014;37(3):475–82.
- [14] Kouloulis VE, Poortmans PM, Bernier J, Horiot JC, Johansson KA, Davis B, et al. The Quality Assurance programme of the radiotherapy group or European Organisations for Research and Treatment of Cancer (EORTC): a critical appraisal of 20 years of continuous efforts. *Eur J Cancer* 2003;29(4):430–7.
- [15] Ibbott GS, Followill DS, Molineu HA, Lowenstein JR, Alvarez PE, Roll JE. Challenges in Credentialing Institutions and Participants in advanced technology Multi-institutional Clinical trials. *Int J Radiat Oncol Biol Phys* 2008;71:S71–5.
- [16] Budiharto T, Musat E, Poortmans P, Hurkmans C, Monti A, Bar-Deroma R, et al. Profile of European radiotherapy departments contributing to the EORTC Radiation Oncology Group (ROG) in the 21st century. *Radiother Oncol* 2008;88(3):403–10.
- [17] McGarry CK, Agnew CE, Hussein M, Tsang Y, McWilliam A, Hounsell AR, et al. The role of complexity metrics in a multi-institutional dosimetry audit of VMAT. *Br J Radiol* 2016;1057(89).
- [18] Ezzell GA, Burmeister JW, Dogan N, LoSasso TJ, Mechalakos JG, Mihailidis D, et al. IMRT commissioning: Multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119. *Med Phys* 2009;36(11):5359–73.
- [19] Svenssons E. Evaluation of complexity and deliverability of IMRT-treatment plans (Master thesis). University of Gothenburg; 2011.
- [20] Jurado-Bruggeman D, Hernández V, Sáez J, Navarro D, Pino F, Martínez T, et al. Multi-centre audit of VMAT planning and pre-treatment verification. *Radiother Oncol* 2017;124(2):302–10.
- [21] Hernandez V, Abella R, Calvo JF, Jurado-Bruggemann D, Sancho I, Carrasco P. Determination of the optimal tolerance for MLC positioning in sliding window and VMAT techniques. *Med Phys* 2015;4:1911–6.
- [22] Storey JD. A direct approach to false discovery rates. *J R Stat Soc Ser B* 2002;64:479–98.
- [23] Storey JD, Bass AJ, Dabney A, Robinson D. *qvalue: Q-value estimation for false discovery rate control*. R Package 2.6.0. 2015. <http://github.com/jdstorey/qvalue>.
- [24] Younge KC, Roberts D, Janes LA, Anderson C, Moran JM, Matuszak MM. Predicting deliverability of volumetric-modulated arc therapy (VMAT) plans using aperture complexity analysis. *J Appl Clin Med Phys* 2016;17(4):124–31.
- [25] Webb S. Use of a quantitative index of beam modulation to characterize dose conformality: illustration by a comparison of full beamlet IMRT, few-segment IMRT (fsIMRT) and conformal unmodulated radiotherapy. *Phys Med Biol* 2003;48(14):2051–62.
- [26] McGarry CK, Chinneck CD, O’Toole MM, O’Sullivan JM, Prise KM, Hounsell AR. Assessing software upgrades, plan properties and patient geometry using intensity modulated radiation therapy (IMRT) complexity metrics. *Med Phys* 2011;38(4):2027–34.