



ELSEVIER

Contents lists available at ScienceDirect

Physica Medica

journal homepage: www.elsevier.com/locate/ejmp

Original paper

Use of metrics to quantify IMRT and VMAT treatment plan complexity: A systematic review and perspectives



Mikaël Antoine^{a,b,*}, Flavien Ralite^{b,c}, Charles Soustiel^d, Thomas Marsac^b, Paul Sargos^b, Audrey Cugny^b, Jérôme Caron^b

^a Service d'onco-radiothérapie, Polyclinique de Bordeaux Nord, 33000 Bordeaux, France

^b Department of Radiotherapy, Institut Bergonié, Comprehensive Cancer Centre, F-33000 Bordeaux, France

^c SUBATECH, IMT-Atlantique, CNRS/IN2P3, Université de Nantes, Nantes, France

^d Department of Radiotherapy, Centre Hospitalier de Dax, Dax, France

ARTICLE INFO

Keywords:

Modulation indices

Plan complexity

Volumetric modulated arc therapy

ABSTRACT

Purpose: Fixed-field intensity modulated radiation therapy (FF-IMRT) or volumetric modulated arc therapy (VMAT) beams complexity is due to fluence fluctuation. Pre-treatment Quality Assurance (PTQA) failure could be linked to it. Several plan complexity metrics (PCM) have been published to quantify this complexity but in a heterogeneous formalism. This review proposes to gather different PCM and to discuss their eventual PTQA failure identifier abilities.

Methods and materials: A systematic literature search and outcome extraction from MEDLINE/PubMed (National Center for Biotechnology Information, NCBI) was performed. First, a list and a synthesis of available PCM is made in a homogeneous formalism. Second, main results relying on the link between PCM and PTQA results but also on other uses are listed.

Results: A total of 163 studies were identified and $n = 19$ were selected after inclusion and exclusion criteria application. Difference is made between fluence and degree of freedom (DOF)-based PCM. Results about the PCM potential as PTQA failure identifier are described and synthesized. Others uses are also found in quality, big data, machine learning and audit procedure.

Conclusions: A state of the art is made thanks to this homogeneous PCM classification. For now, PCM should be seen as a planning procedure quality indicator although PTQA failure identifier results are mitigated. However limited clinical use seems possible for some cases. Yet, addressing the general PTQA failure prediction case could be possible with the big data or machine learning help.

1. Introduction

Fixed field-intensity modulated radiation therapy (FF-IMRT) and volumetric modulated arc therapy (VMAT) have become common in radiation oncology treatments of gynaecological, prostate, and head and neck (H&N) tumours [1]. Compared to older techniques, the estimated advantages are dose conformation improvement, dose escalation potential, simultaneous integrated boost feasibility or highest organ at risk sparing performance [2–6]. At this time, publications relating to commissioning [7], treatment planning [8], associated quality assurance (QA) [9], clinical implementation [10,11], dose prescription and reporting [12] have been emitted, making this technique widespread.

In association with this technique a dedicated QA program [13,14] is needed, and a distinction between linac QA (LQA), pre-treatment QA

(PTQA) and patient-specific QA (PSQA) could be made. Linac QA is related to the linac capacities of conducting accurate FF-IRMT or VMAT beams delivery by, for example, realizing specific multi-leaf collimator (MLC) tests [15,16]. Performed before the start of treatment, PTQA relies on the criteria-based validation of beam delivery detector acquisition. Last, PSQA is related to an in-vivo measurement made during treatment with detectors such as diodes or through EPID-transit solutions.

A report was recently published on methodologies and tolerance limits [17] permitting the harmonisation of PTQA practices and establishing consistent and comparable criteria among institutions. As the origins of PTQA failure could be uneasy to identify between dose calculation (TPS, dedicated software), dose delivery (linac) or dose measurement (detectors) [18], most of the time re-planning is necessary.

* Corresponding author at: Service d'oncoradiothérapie, Polyclinique de Bordeaux Nord, 15, rue Claude-Boucher, 33000 Bordeaux, France.

E-mail address: m.antoine@bordeauxnord.com (M. Antoine).

<https://doi.org/10.1016/j.ejmp.2019.05.024>

Received 5 February 2019; Received in revised form 24 May 2019; Accepted 26 May 2019

Available online 09 July 2019

1120-1797/ © 2019 Associazione Italiana di Fisica Medica. Published by Elsevier Ltd. All rights reserved.

Moreover, PTQA – and specifically the gamma pass rate (γ_{pass}) index [19,20] – correlates weakly with dose-volume histogram (DVH) variation [21,22] and may not substantially detect clinically relevant errors [23]. Pre-treatment QA failure identifiers could then be considered as less time-consuming than linac-occupation approaches and would represent a benefit for radiation oncology departments.

So, a VMAT or FF-IMRT plan could be considered as the simultaneous variation of different degrees of freedom (DOFs) combined with fixed-physical properties (FPPs). A DOF corresponds to all achievable beam discrete values which varies during deliverance like dose rate, gantry rotation speed and monitor units (MU) per control point (CP) [24], and FPPs to constant parameters such as MLC properties (for example leaves transmission, edge shape and motion limitations), linac properties (such as beam energy and maximum gantry rotation speed) or plan specificities (for instance CP number). The beam fluence depends on both DOF and FPP combinations, and its fluctuation may be seen as its complexity.

Variables could quantify DOF variations such as the overall leaf travel (OLT), which represents the total distance (in mm) travelled by an MLC leaf during beam delivery. Given the simple variable employment, OLT is a basic DOF calculation; however, to go further, some authors have proposed plan complexity metrics (PCMs). These metrics are DOF and FPP-based calculation results and could be seen as beam complexity quantification. One of the potential uses of PCMs could be their PTQA failure prediction capacities.

To date, a substantial amount of literature exists since one of the earliest contributions, which was made by the Webb team with the introduction of the modulation index (MI) in 2003 [25]. Giorgia [26] continued with this approach with the first MI-PTQA failure correlation tests. In 2010, McNiven [27] introduced the modulation complexity score (MCS) as a different PCM combination. Yet, many other authors have contributed to this research field, and PCMs were heterogeneously formalised. A semantic issue also appeared because words such as metrics, modulation indices or plan complexity score are becoming common and often refer to the same concept. Therefore, a systematic review, such as the one presented here, should constitute an interesting approach not only to provide a homogeneous overview of already published PCMs and to discuss their known PTQA failure prediction abilities, but also to address their other potential uses.

2. Material and methods

According to preferred reporting items for systematic reviews and meta-analyses (PRISMA)'s recommendations for systematic review [28], published articles were identified in April 2019 through a Medline PubMed search using a combination of ('VMAT' OR 'IMRT') AND ('complexity' OR 'modulation') AND ('index' OR 'metric' OR 'metrics' OR 'indices' OR 'score'). No date range was applied, and only English-written and peer-reviewed articles were considered. Inclusion criteria were as follows: articles with new DOF or fluence map-related PCM explicit formulas or articles involving treatment with PCM adaptation from existing metrics. Exclusion criteria were review and optimisation function modification articles. First, published PCMs are reviewed with the following aims: to class, synthesise and homogenise the formalism and to qualitatively describe the influence parameters (1). Second, the main results dealing with the PCM PTQA-failure prediction abilities are described (2). Third, PCMs' other uses are presented (3).

3. Results

The PubMed search yielded 163 results, of which 159 are original, while 4 are review articles. After applying both inclusion and exclusion criteria, only 19 articles were selected [15,25–27,29–44].

3.1. Review of existing plan complexity metrics

The classification, synthesis and qualitative evaluation of already published PCMs can be differentiated according the metrics' relate to a beam fluence map (MI, MI_G, PIMV, FMC, FD, ASM, IDM, CTR, VAR, COR, S [Table 1]) or to a DOF-based calculation (MI_s, MI_a, MI_{SPORT}, DC, M, MCS, MCS_v, oMCS, MFA, CAM, EAM, C/A, SAS, MAD, CLS, CAS, PA, PI, PM, PMU, MLC_{velo}, ALPO [Table 2]).

3.2. Main results regarding the link between PCMs and pre-treatment QA failure

Results are synthesised in Table 3. A total of 23 correlation tests were found in the literature. With the exception of Götstedt [42], none of them found a significant and strong correlation between PCM and PTQA results, but such a correlation was obtained for MLC complexity tests. Despite this, some other results were also found.

For FF-IMRT, after the Webb MI [25] adaptation, Giorgia [26] realised some correlation tests between MI_G and EPID $\gamma_{3\%/3mm}$ and argued in favour of fixing an MI_G threshold at 19. Park [38] added three other PCMs (MI_s, MI_a, MI_t) to the MI family, arguing that they focus on mechanical parameter variations under the assumption that abrupt mechanical variations increase delivery uncertainties. A correlation analysis was performed between MI_s, MI_a, MI_t, MCS_v, LTMCS and MI_{SPORT} and γ_{pass} rates (3%/3 mm, 2%/2 mm, 1%/2 mm and 2%/1 mm) from 2D diode array forty planar dose distribution acquisitions. The significant obtained Rs values for MI_s, MI_a and MI_t with 2%/2 mm criteria were -0.637 , -0.648 and -0.660 respectively, and with 1%/2 mm criteria, those values were -0.662 , -0.668 and -0.675 respectively. It should be noted that the MI_s values were greater for H&N than for prostate, which is a result also found by Li [32], confirming intuitive thought on the greater complexities of H&N plans.

McNiven tested the link between different localisations of the FF-IMRT MCS and $\gamma_{3\%/3mm}$ and $\gamma_{2\%/1mm}$ pass rates acquired on a 2D diode array [27]. First, it was demonstrated that MU and MU/CP are weakly correlated with MCS because of the higher amount of information contained in MCS compared to intuitive basic PCMs such as MU or MU/CP. Second, results indicated that MCS could have a threshold effect, and beams higher than 0.8 MCS were identified as robust (100% specificity). Masi [33] completed this work by adapting MCS for a VMAT plan (MCS_v) and particularly studied the CP angular sampling impact. A Pearson correlation test was performed between LT, MCS_v and LTMCS_v and local $\gamma_{3\%/3mm}$ and $\gamma_{2\%/2mm}$ pass rates on a bi-planar diode array. A significant correlation was found for 4° CP sampling, and those PCMs were found to be possible PTQA failure identifiers while using threshold values. It was also found that a finer angular sampling increased γ pass rate results but lowered the correlation.

Young [45] introduced the M PCM and performed a receiver operating characteristic (ROC) analysis for 649 previously treated plans completed by 62 plans for which PTQA failed. This allowed for the selection of a threshold value of 0.180 mm^{-1} , which led to the true positive rate of 44% for correctly identifying PTQA failed plans and to the false positive rate of 7%. According to these results, the implementation of overall plan screening was set up in the clinical workflow, and plans that were too complex were excluded because of their PTQA-failure risk.

Valdes [46] tested 78 metrics on 498 IMRT plans and suggested that $\gamma_{3\%/3mm}$ local dose/DTA and 90% threshold PTQA failures have five origins: MLC leaves transmission, leaf end leakage, jaw transmission, tongue and groove effect, and charge particle disequilibrium. It was found that the most relevant metrics to describe the passing rates were the MU factor (MU per Gy), the small aperture score (SAS), the irregularity factor and the fraction of the plan delivered at the corners of

Table 1
Review of IMRT/VMAT plan Plan Complexity Metrics formulas calculated on beams fluence map.

PCM	Type PCM Name Reference	Formula	Plan Complexity Sensitivity
Modulation Index	Original MI Webb, <i>Phys. Med. Biol.</i> , 2003 Adaptation MI _G Giorgia, <i>Radiat. Oncol.</i> , 2007	$MI_I = \int_0^F Z(f)df$ <p>where $Z_I(f) = Z_{I_x}(f) = \frac{1}{n-1}N_{I_x}(f; \Delta I_x > f\sigma_I)$</p> $MI_G = \int_0^F Z_I(f)df$ <p>where $Z_I(f) = \frac{ Z_{I_x}(f) + Z_{I_y}(f) + Z_{I_{xy}}(f) }{3}$</p> $Z_{I_x}(f) = \frac{1}{n(m-1)}N_{I_x}(f; \Delta I_x > f\sigma_I);$ $Z_{I_y}(f) = \frac{1}{(n-1)m}N_{I_y}(f; \Delta I_y > f\sigma_I); Z_{I_{xy}}(f) = \frac{1}{(n-1)(m-1)}N_{I_{xy}}(f; \Delta I_{xy} > f\sigma_I)$	Sensitive to the intensity fluence map in one direction, (take account of intensity values changes between adjacent bixels of the fluence map in X direction). Sensitive to the intensity fluence map in three directions, (take account of intensity values changes between adjacent bixels of the fluence map in X, Y and XY directions).
Plan Intensity Map Variation	Original PIMV Coselmon, <i>Med. Phys.</i> , 2005.	$PIMV = \sum_{i=1}^n \sum_{j=1}^m (I_{i,j} - I_{i,j+1} + I_{i,j} - I_{i+1,j} + I_{i,j} - I_{i+1,j+1})$	Sensitive to the intensity fluence map in two directions (X and Y).
Fluence Map Complexity	Original FMC Llacer, <i>Phys. Med. Biol.</i> , 2001	$FMC = \frac{1}{\sum_{i=1}^n l_i} \sqrt{\sum_{i=1}^n (l_i - \lambda[l_{i+1} + l_{i-1}])^2}$	Sensitive to the intensity fluence map in one direction by considering the difference between the fluence of a bixel and the fluence of lateral neighboring bixels.
Edge Area Metric	Original EAM Götstedt, <i>Med. Phys.</i> , 2015	$EAM = \overline{EAM}_k$ <p>where</p> $EAM_k = \frac{R_{edge}}{R_{edge} + R_{openarea}}$	Sensitive to the relative amount of edge region for the MLC aperture. Note that the EAM of a beam is the mean value of all EAM scores calculated for each CP.
Fractal Dimension	Original FD M. Nauta, <i>Med. Phys.</i> , 2011. M. Tambasco, <i>Phys. Medica</i> , 2013.	$FD = \frac{4 - SLP}{2}$ <p>where SLP is the slope of the plot of $\log(\gamma(h))$ versus $\log(h)$ With</p> $\gamma(h) = c \cdot h^{4-2 \cdot FD} = \frac{1}{2D} \sum_{i=1}^D (FS(x_i) - FS(x_i + h))^2$	Sensitive to the structural irregularities of the fluence map at different size scale, by considering the fractal surface variations between neighbor pixels, spaced of a distance h. Note that the mathematical expression is derived from the variogram method.
Angular Second Moment	Original ASM Park, <i>Radiat. Oncol. Lond. Engl.</i> , 2014.	$ASM = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} GLCM_{ij}^2$	Metric based on the GLCM that indicates a measure of the homogeneity of a fluence map.
Inverse Difference Moment	Original IDM Park, <i>Radiat. Oncol. Lond. Engl.</i> , 2014.	$IDM = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{1}{1 + GLCM_i - GLCM_j } GLCM_{ij}$	Metric based on the GLCM that indicates a measure of the local homogeneity of a fluence map.
Contrast	Original CTR Park, <i>Radiat. Oncol. Lond. Engl.</i> , 2014.	$CTR = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} GLCM_i - GLCM_j ^2 \cdot GLCM_{ij}$	Metric based on the GLCM that indicates a measure of the local variations in a fluence map.
Variance	Original VAR Park, <i>Radiat. Oncol. Lond. Engl.</i> , 2014.	$VAR = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (GLCM_i - \overline{GLCM}_i)^2 \cdot GLCM_{ij} + \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (GLCM_j - \overline{GLCM}_j)^2 \cdot GLCM_{ij}$	Metric based on the GLCM that indicates a measure of the inhomogeneity of a fluence map.
Correlation	Original COR Park, <i>Radiat. Oncol. Lond. Engl.</i> , 2014.	$COR = \frac{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (GLCM_i - \overline{GLCM}_i) \cdot (GLCM_j - \overline{GLCM}_j) \cdot GLCM_{ij}}{\sigma_x \sigma_y}$	Metric based on the GLCM that indicates a measure of the linear dependency of gray levels in a fluence map.
Entropy	Original S Park, <i>Radiat. Oncol. Lond. Engl.</i> , 2014.	$S = - \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} GLCM_{ij} \cdot \log(GLCM_{ij})$	Metrics based on the GLCM that indicates a measure of a randomness of a fluence map.

the 40-cm × 40-cm field. Indeed, according to Valdes, the higher these values were, the lower the PTQA passing rate was.

Park [47] tested the correlation of MIs, MIa, MIc, MCS, PA and PI for 202 failure occurrences of FF-IMRT with PTQA on two different

linacs and with both Mapcheck2 and Archeck. The author found that PI was the best PTQA failure identifier and concluded by arguing that ‘the PI value could support the verification of IMRT plan delivery accuracies before patient treatment and reduce resource consumption in the clinic, as it

Table 2
Review of IMRT/VMAT plan Plan Complexity Metrics formulas based on Degrees Of Freedom variation.

PCM	Type PCM Name Reference	Formula	Plan Complexity Sensitivity (Influence factors)
Modulation Index	Adaptation MI _s Park Phys. Med. Biol. 2014	$MI_s = \sum_{l=1}^{N_{leaf}} MI_{MLCs_l}$ $MI_{MLCs_l} = \int_0^F Z_{MLCs_k,l}(f) df$ <p>where $Z_{MLCs_k,l} = \frac{1}{N_{CP}-1} N_{MLCs_k,l}(f; MLCs_{k,l} > f\sigma_{MLCs_k,l})$</p>	MLC leaves speed between different control point (CP) (all leaves positions in each CP and the time of each CP).
	Adaptation MI _a Park Phys. Med. Biol. 2015	$MI_a = \sum_{l=1}^{N_{leaf}} MI_{MLCa_l}$ <p>with</p> $MI_{MLCa_l} = \int_0^F Z_{MLCa_k,l}(f) df$ <p>where</p> $Z_{MLCa_k,l} = \frac{1}{N_{CP}-2} N_{MLCa_k,l}(f; MLCa_{k,l} > \alpha f\sigma_{MLCa_k,l})$	MLC leaves speed and acceleration (all leaves positions in each CP and the time of each CP weighted with the ponderation time factor).
	Adaptation MI _t Park Phys. Med. Biol. 2015	$MI_t = \sum_{l=1}^{N_{leaf}} MI_{MLCt_l}$ <p>where</p> $MI_{MLCt_l} = \int_0^K Z_{total}(f) df$ <p>K = 0.2,0.5,1,2 where</p> $Z_{total}(f) = \left(\frac{1}{N_{CP}-2} \right) \sum_{i=1}^{N_{CP}} \{N_i(f; MLC \text{ speed}_i > f\sigma_{MLC \text{ speed}}). W_{GA,i+1} W_{MU,i+1}\}$	MI evaluating the speed of MLC, acceleration of MLC, gantry rotation acceleration and DR variation comprehensive
	Adaptation MI _{SPORT} Li & Xing, Med. Phys. 2013.	$MI_{SPORT} = \sum_{k=1}^{-K:K} \left[\sum_{l=1}^{N_{leaf}} (p_{l,left \text{ bank},k} - p_{l,left \text{ bank},k+1} + p_{l,right \text{ bank},k} - p_{l,right \text{ bank},k+1}) \right] \cdot \left \frac{MU_k - MU_{k+1}}{GA_k - GA_{k+1}} \right $	Sensitive to the level of intensity modulation of a CP, by considering the leaves positions of the MLC, MU and Gantry Angle (GA) at the CP considered and for the -Kth and Kth CPs neighboring the CP considered. FF-IMRT: Sensitive to the Monitor Units (MU) and to the total number of segments of an IMRT plan.
Delivery Complexity	Original DC Anker, J. Appl. Clin. Med. Phys, 2010	$DC = \frac{MU \cdot N_{segment,plan}}{MU \cdot N_{segment,default \ plan}}$	Sensitive to field aperture per CP, by considering the ratio between the MLC aperture perimeter and the area weighted by Monitor Units (MU) per CP.
Index of Modulation	Original M Young, Med. Phys 2012	$M = \frac{1}{MU} \sum_{k=1}^{N_{CP}} MU_k \cdot \frac{P_k}{A_k}$	Sensitive to the aperture area variability (AAV) and the leaf sequence variability (LSV) per segment for an IMRT beam.
Modulation Complexity Score	Original MCS McNiven, Med. Phys., 2010	$MCS = \sum_{k=1}^{N_{segment}} AAV_k \cdot LSV_k \cdot \frac{MU_k}{MU}$ <p>where</p> $AAV_k = \left(\frac{\sum_{l=1}^{N_{leaf}} (p_{l,left \text{ bank}} - p_{l,right \text{ bank}})}{\sum_{l=1}^{N_{leaf}} (\max(p_{l,left \text{ bank}}) - \max(p_{l,right \text{ bank}}))} \right)_k$ $LSV_k = \left(\frac{\sum_{l=1}^{N_{leaf}} [p_{max} - (p_l - p_{l+1})]}{N_{leaf} \cdot P_{max}} \right)_{left \text{ bank},k} \left(\frac{\sum_{l=1}^{N_{leaf}} [p_{max} - (p_l - p_{l+1})]}{N_{leaf} \cdot P_{max}} \right)_{right \text{ bank},k}$	VMAT. Sensitive to the aperture area variability (AAV) and the leaf sequence variability (LSV) per CP.
	Adapted MCS _v Masi, Med. Phys., 2013.	$MCS_v = \sum_{k=1}^{N_{CP}} \left(\frac{AAV_k + AAV_{k+1}}{2} \cdot \frac{LSV_k + LSV_{k+1}}{2} \right) \cdot \frac{MU_k}{MU}$ <p>where</p> $AAV_k = \left(\frac{\sum_{l=1}^{N_{leaf}} (p_{l,left \text{ bank}} - p_{l,right \text{ bank}})}{\sum_{l=1}^{N_{leaf}} (\max(p_{l,left \text{ bank}}) - \max(p_{l,right \text{ bank}}))} \right)_k$ $LSV_k = \left(\frac{\sum_{l=1}^{N_{leaf}} [p_{max} - (p_l - p_{l+1})]}{(N_{leaf}-1) \cdot P_{max}} \right)_{left \text{ bank},k} \left(\frac{\sum_{l=1}^{N_{leaf}} [p_{max} - (p_l - p_{l+1})]}{(N_{leaf}-1) \cdot P_{max}} \right)_{right \text{ bank},k}$	Sensitive to the aperture area variability (AAV) and to the sector leaf sequence variability (sLSV), which considers a specific organ located in the field, per segment for an IMRT beam.
Modulation Complexity Score	Adapted oMCS Sumida, J. Radiat. Res., 2017.	$oMCS = \sum_{k=1}^{N_{segment}} AAV_k \cdot sLSV_k \cdot \frac{MU_k}{MU}$ <p>where</p> $sLSV_k = \left(\frac{\sum_{l=1}^{N_{leaf \ organ}} [p_{max} - (p_l - p_{l+1})]}{(N_{leaf}-1) \cdot P_{max}} \right)_{left \text{ bank},k} \left(\frac{\sum_{l=1}^{N_{leaf \ organ}} [p_{max} - (p_l - p_{l+1})]}{(N_{leaf}-1) \cdot P_{max}} \right)_{right \text{ bank},k}$	Sensitive to the aperture area variability (AAV) and to the sector leaf sequence variability (sLSV), which considers a specific organ located in the field, per segment for an IMRT beam.

(continued on next page)

Table 2 (continued)

PCM	Type PCM Name Reference	Formula	Plan Complexity Sensitivity (Influence factors)
Mean Field Area	Original MFA Crowe, <i>Australas. Phys. Eng. Sci. Med.</i> , 2014	$MFA = \sum_{k=1}^{N_{CP}} A_k \cdot \frac{MU_k}{MU}$	Sensitive to field aperture per CP, by considering only the MLC aperture area weighted by the Monitor Units (MU) per CPs.
Converted Aperture Metric	Original CAM Götstedt, <i>Med. Phys.</i> , 2015	$CAM = \bar{CAM}_k$ where $CAM_k = 1 - g(d_i)_k \cdot g(A_{Eq})_k$ $g(x) = 1 - e^{-x}$	Sensitive to the field aperture per CP, by considering distance between the MLC leaves in both X and Y directions. Note that the CAM of a beam is the mean value of all CAM scores calculated for each CP.
Edge Area Metric	Original EAM Götstedt, <i>Med. Phys.</i> , 2015	$EAM = \bar{EAM}_k$ where $EAM_k = \frac{R_{edge}}{R_{edge} + R_{openarea}}$	Sensitive to the relative amount of edge region for the MLC aperture. Note that the EAM of a beam is the mean value of all EAM scores calculated for each CP.
Circumference/area	Original C/A Götstedt, <i>Med. Phys.</i> , 2015	$C/A = \sum_{k=1}^{N_{CP}} \frac{P_k}{A_k}$	Sensitive to the field aperture per CP, by considering the ratio between the MLC aperture perimeter (or circumference) and the area.
Small Aperture Score	Original SAS Crowe, <i>Australas. Phys. Eng. Sci. Med.</i> , 2014	$SAS = \sum_{k=1}^{N_{CP}} \frac{N_{leaf}(a_c > a_l > 0)_k}{N_{leaf}(a_l > 0)_k} \cdot \frac{MU_k}{MU}$	Sensitive to the aperture per CP by considering the distance between opposite leaves under a certain criteria.
Mean Asymmetry Distance	Original MAD Crowe, <i>Australas. Phys. Eng. Sci. Med.</i> , 2014	$MAD = \sum_{k=1}^{N_{CP}} \left(\sum_{l=1}^{N_{leaf}} C_{leaf} - C_{MLC} \right) \cdot \frac{MU_k}{MU}$	Sensitive to the aperture per CP by considering the average of the distance between the centre of the aperture distance between opposite leaf pairs and the MLC central axis.
Closed Leaf Score	Original CLS Crowe, <i>Australas. Phys. Eng. Sci. Med.</i> , 2014	$CLS = \sum_{k=1}^{N_{CP}} \frac{N_{leaf}(a_l > 0)_k}{N_{leaf,k}} \cdot \frac{MU_k}{MU}$	Sensitive to the aperture per CP by considering the closed leaves.
Cross-Axis Score	Original CAS Crowe, <i>Australas. Phys. Eng. Sci. Med.</i> , 2014	$CAS = \sum_{k=1}^{N_{CP}} \frac{N_{leaf}(a_l > C_{MLC})_k}{N_{leaf}(a_l > 0)_k} \cdot \frac{MU_k}{MU}$	Sensitive to the aperture per CP by considering the leaves that cross the MLC central axis.
Plan Average Beam Area (PA)	Original PA Du, <i>Med. Phys.</i> , 2014.	$PA = \frac{\sum_{k=1}^{beam} BA_i \cdot MU_i}{MU_p}$ where $BA_i = \frac{\sum_{j=1}^{segment} MU_{ij} AA_{ij}}{MU_i}$ where $AA_{ij} = \sum_{k=1}^{N_{leafpair}} t_k \cdot (2_{ijk} - 1_{ijk})$	Sensitive to field aperture per CP, by considering only the MLC aperture area weighted by Monitor Units (MU) per CP.
Plan Average Beam Irregularity (PI)	Original PI Du, <i>Med. Phys.</i> , 2014.	$PI = \frac{\sum_{k=1}^{beam} BI_i \cdot MU_i}{MU_p}$ where $BI_i = \frac{\sum_{j=1}^{segment} MU_{ij} AI_{ij}}{MU_i}$ where $AI_{ij} = \frac{AP^2_{ij}}{4\pi AA_{ij}}$	Sensitive to the field aperture irregularity per CP, by considering the non-circularity of the aperture area.

(continued on next page)

Table 2 (continued)

PCM	Type PCM Name Reference	Formula	Plan Complexity Sensitivity (Influence factors)
Plan Averaged Beam Modulation (PM)	Original PM Du, <i>Med. Phys.</i> , 2014.	$PM = 1 - \frac{\sum_{k=1}^{beam} BM_i \cdot MU_i}{MU_p}$ where $BM_i = \frac{\sum_{j=1}^{segment} MU_{ij} A_{ij}}{MU_i \cdot U(A_{ij})}$ U(A _{ij}) is the union area of all apertures of beam i.	Sensitive to the field aperture area per CP by considering the union area of all aperture areas of a beam.
Plan Normalized MU (PMU)	Original PMU Du, <i>Med. Phys.</i> , 2014.	$PMU = \frac{MU_p \cdot 2Gy}{d}$ With d the prescribed dose per fraction (Gy)	
MLC leaf velocity	Original MLC _{velo} Agnew, <i>J. Appl. Clin. Med. Phys.</i> , 2014.	$MLC_{velo} = \frac{pl_k - pl_{k+1}}{t_k}$	Sensitive to the mechanical delivery inaccuracies of the MLC at each CP, by considering the ratio between the distance travelled by an active MLC leaf between two consecutive CPs and the time between two consecutive CPs.
Average Leaf Pair Opening (ALPO)	Original ALPO Zygmanski, <i>Med. Phys.</i> , 2001.	$ALPO = \frac{\sum_{k=1}^{N_{CP}} \sum_{l=1}^{N_{leaf}} pl_{l, right\ bank} - pl_{l, left\ bank} \cdot MU_k}{\sum_{k=1}^{N_{CP}} \sum_{l=1}^{N_{leaf}} MU_k (for\ a_l \neq 0)}$	Sensitive to MLC gap error by considering the ratio between the sum of the aperture area and the sum of the fractional MU during which a leaf pair is open.

can be calculated at the planning level.’ Lastly, Glenn [48] studied the relationship between 16 PCMs (including MU, MCS, PI, MI_s, MI_a and MI) on H&N plans and found no significant correlation between PCM results (r_s = ±0.30) and γ_{7%/4mm} radio-chromic films measurements or single-point thermo-luminescent dosimeter (TLD) dose.

3.3. Other PCM uses

3.3.1. Data statistics and machine learning

Palaniswaamy [49] developed in-house software to monitor the statistical trends of PTQA results differentiating radiotherapy localisations. By setting the specific-site PTQA tolerance, false positive and false negative results were reduced. Valdes [46] pooled over 78 variables, which included PCMs and FPPs (such as dosimetric leaf gap, leaf transmission and SAS), and applied a self-developed machine-learning algorithm to predict the PTQA γ_{pass} rate. This could be seen as a virtual PTQA, as the team was able to realise an *a posteriori* predictability of PTQA results with a 3% confidence level.

3.3.2. Audit perspectives and comparison of centres

Nelms [50] used a DVH-based PCM set to quantify clinical practice variations and potential technology parameter dependence. It was demonstrated that a user’s skill seems to be more important than technology or demographic user characteristics (such as years of experience). McGarry [51] compared 39 VMAT γ_{pass} rates obtained from 34 different centres, as these plans were created from type 1 or type 2 TPS. A significant finding (p < 0.01) was that type 2 TPS created poorer plan quality (higher MCS and MU) and that type 2 had a lower gamma pass rate than type 1 TPS plans when comparing them on a multi-detector.

Hernandez [52] developed an in-house program to generate plans with different complexity indices and compared 100 VMAT plans from two institutions. As stated, it was not possible to use PCMs as PTQA failure identifiers. It was also found that some PCMs addressed the same information, as they are correlated, thereby leading the author to plead for their careful use in multi-centre comparison.

4. Discussion

4.1. Review of existing PCMs

Quality has become one of the major interests of radiotherapy teams [53], especially for patient safety [54]. While medical physics work is increasingly being divided between physicists, dosimetrists, nurses, technicians and medical physicists, clinical involvement should remain a strong reality, especially in QA and risk assessment [55]. To this end, PCM use in a clinical workflow could offer three relevant tools. First, for each FF-IMRT or VMAT plan, a PCM comparison with other similar plans (such as localisation or Linac) would permit one to upgrade the safety of the planning procedure by setting limits and producing knowledge on ‘abnormal plans’. Second, PCM use offers a ‘common language’ between team members, allowing for scientifically based discussions about many topics such as PTQA failure, planning procedures and education. Third, PCM plan quantification should be seen as an evidence production, making it an indicator of the quality of clinical practice. All of this could also be seen as contributing to the confidence of the radiation therapy team.

4.2. Main results regarding the link between PCMs and PTQA failure

Careful consideration is required when addressing the PTQA γ_{pass} problem and one should be aware of what is relevant in PTQA failure. This was well described by Crowe [56], who investigated the action level and PTQA-device dependence and concluded that the use of a γ_{3%/3mm} score for PTQA is widespread despite concerns about the suitability of the γ evaluation, and who suggests that a 3%/3 mm is not sensitive enough. Moreover, Nelms [57] argues for the retirement of this criteria because of its inability to detect systematic errors. It is to say that every correlation test and predictability remain dependent on detector characteristics [58] and their weak capacities to accurately distinguish acceptable from unacceptable plans [59]. Further evaluations must be considered with this PTQA sensitivity problem. This review does not establish a clear general correlation between PCM and PTQA performance and failure. Indeed, while most of the results

Table 3
Main results about correlation tests between Pre-Treatment Quality Assurance score and Plan Complexity Metrics.

PCM Name Reference	Results Reference Technical environment (Delivering mode and Linac ; TPS ; optimization and calculation dose algorithm, measurement device)	Localizations, Patient Number, Type of test	PCM-PTQA link results																		
				2%/2mm	2%/1mm	1%/2mm															
MI _S Park Phys. Med. Biol. 2015	Park Phys. Med. Biol. 2014	<ul style="list-style-type: none"> Prostate + Head and Neck N=40 (20each) Spearman's rho 	Rs (p value)	-0.637 (< 0.001)	-0.471 (0.002)	-0.657 (< 0.001)															
MI _a Park Phys. Med. Biol. 2015	<ul style="list-style-type: none"> VMAT Varian Trilogy with MLC Millenium + TrueBeam STx; MLC HD; PRO3 v10, AAAv10, Elipse; calc grid 2.5mm; 		Rs (p value)	-0.663 (< 0.001)	-0.561 (< 0.001)	-0.669 (< 0.001)															
MI _t Park Phys. Med. Biol. 2015	NB : MIs (f=1) Webb formalism		Rs (p value)	-0.667 (< 0.001)	-0.552 (< 0.001)	-0.669 (< 0.001)															
MI _{SPORT} Li & Xing, Med. Phys., 2013.	Park Phys. Med. Biol. 2014 VMAT, Varian Trilogy with MLC Millenium and TrueBeam STx MapCHECK2 NB : MIs (f=1) Webb formalism MLC HD; PRO3 v10, AAAv10, Elipse; calc grid 2.5mm		Rs (p value)	-0.455 (0.003)	-0.49 (0.001)	-0.502 (0.001)															
MI Crowe, Australas. Phys. Eng. Sci. Med., 2014	Crowe, Australas. Phys. Eng. Sci. Med., 2014 FF-IMRT, TPS Brainlab, MapCHECK2	<ul style="list-style-type: none"> Prostate N=122 beams F-value linear relationship test (p-value) 	<table border="1"> <tr> <td></td> <td colspan="2">2%/2mm</td> </tr> <tr> <td>F (p-value)</td> <td colspan="2">11.397 (0.001)</td> </tr> </table>				2%/2mm		F (p-value)	11.397 (0.001)											
	2%/2mm																				
F (p-value)	11.397 (0.001)																				
MCS McNiven, Med. Phys., 2010	<ul style="list-style-type: none"> Gösted, Med. Phys., 2015 FF-IMRT and VMAT; Varian Clinac iX; Eelipse; AAA; calc grid 2.5mm; Varian EPID aSi 1000 and Gafchromic™ EBT3 film 	<ul style="list-style-type: none"> Various MLC openings N=30 Pearson's correlation between dose difference and PCM value 	<table border="1"> <tr> <td></td> <td colspan="2">3% dd</td> <td colspan="2">5% dd</td> </tr> <tr> <td></td> <td>EPID</td> <td>EBT3</td> <td>EPID</td> <td>EBT3</td> </tr> <tr> <td>r</td> <td>0.44</td> <td>0.46</td> <td>0.59</td> <td>0.67</td> </tr> </table>					3% dd		5% dd			EPID	EBT3	EPID	EBT3	r	0.44	0.46	0.59	0.67
	3% dd		5% dd																		
	EPID	EBT3	EPID	EBT3																	
r	0.44	0.46	0.59	0.67																	
MCSv	<ul style="list-style-type: none"> Masi, Med. Phys., 2013. VMAT, Elekta Synergy 1cm MLC; ONCENTRA; Masterplan v4.1 ; PencilBeam algorithm ; DELTA 4 	<ul style="list-style-type: none"> 142 plans different plans Pearson's r correlation 	<table border="1"> <tr> <td></td> <td>3%/3mm</td> <td colspan="2">2%/2mm</td> </tr> <tr> <td>r (4° CP)</td> <td>0.5</td> <td colspan="2">0.54</td> </tr> <tr> <td>r (3°/2° CP)</td> <td>0.48</td> <td colspan="2">0.47</td> </tr> </table>					3%/3mm	2%/2mm		r (4° CP)	0.5	0.54		r (3°/2° CP)	0.48	0.47				
	3%/3mm	2%/2mm																			
r (4° CP)	0.5	0.54																			
r (3°/2° CP)	0.48	0.47																			
	<ul style="list-style-type: none"> Park Phys. Med. Biol. 2015 VMAT, Varian Trilogy with MLC Millenium and TrueBeam STx MapCHECK2; NB : MIs (f=1) Webb formalism; MLC HD; PRO3 v10, AAAv10, Elipse; calc grid 2.5mm 	<ul style="list-style-type: none"> Prostate + Head and Neck N=40 (20each) Spearman's rho 	Rs (p value)	0.186 (0.251)	0.365 (0.021)	0.157 (0.334)															

(continued on next page)

Table 3 (continued)

	Agnew, <i>J. Appl. Clin. Med. Phys.</i> , 2014. Eclipse PRO v10, AAAv10, calc grid 0.25cm, Octavius 4D	VMAT 10H&N, 10 prostate, 10 pelvis Pearson's correlation	<table border="1"> <tr> <td></td> <td>2%/2mm</td> </tr> <tr> <td>Rs (p value)</td> <td>0.654 (<0.001)</td> </tr> </table>		2%/2mm	Rs (p value)	0.654 (<0.001)											
	2%/2mm																	
Rs (p value)	0.654 (<0.001)																	
LTMCS Masi, <i>Med. Phys.</i> , 2013	VMAT, Varian Trilogy with MLC Millenium and TrueBeam STx MapCHECK2 NB : MIs (≠1) Webb formalism MLC HD; PRO3 v10, AAAv10, Elipse; calc grid 2.5mm	<ul style="list-style-type: none"> 142 plans different plans Pearson's r correlation 	<table border="1"> <tr> <td></td> <td>2%/2mm</td> <td>2%/1mm</td> <td>1%/2mm</td> </tr> <tr> <td>Rs (p value)</td> <td>0.312 (0.005)</td> <td>0.371 (0.018)</td> <td>0.343 (0.03)</td> </tr> </table>		2%/2mm	2%/1mm	1%/2mm	Rs (p value)	0.312 (0.005)	0.371 (0.018)	0.343 (0.03)							
	2%/2mm	2%/1mm	1%/2mm															
Rs (p value)	0.312 (0.005)	0.371 (0.018)	0.343 (0.03)															
oMCS Sumida, <i>J. Radiat. Res.</i> , 2017.	FF-IMRT, Siemens ONCOR; TPS XiO; calc grid 2mm; MapCHECK2	<ul style="list-style-type: none"> Head and Neck N=16 Spearman's rho 	<table border="1"> <tr> <td></td> <td>3%/3mm</td> <td>2%/2mm</td> </tr> <tr> <td>r (p value)</td> <td>0.233 (NS)</td> <td>0.403 (NS)</td> </tr> </table>		3%/3mm	2%/2mm	r (p value)	0.233 (NS)	0.403 (NS)									
	3%/3mm	2%/2mm																
r (p value)	0.233 (NS)	0.403 (NS)																
MFA Crowe, <i>Australas. Phys. Eng. Sci. Med.</i> , 2014	FF-IMRT, TPS Brainlab, MapCHECK2	<ul style="list-style-type: none"> Prostate N=122 beams F-value linear relationship test (p-value) 	<table border="1"> <tr> <td></td> <td>2%/2mm</td> </tr> <tr> <td>F (p-value)</td> <td>5.439 (0.021)</td> </tr> </table>		2%/2mm	F (p-value)	5.439 (0.021)											
	2%/2mm																	
F (p-value)	5.439 (0.021)																	
CAM	Götstedt, <i>Med. Phys.</i> , 2015 <ul style="list-style-type: none"> FF-IMRT and VMAT; Varian Clinac iX; Elipse; AAA; calc grid 2.5mm; Varian EPID aSi 1000 and Gafchromic™ EBT3 film 	<ul style="list-style-type: none"> Various MLC openings N=30 Pearson's correlation between dose difference and PCM value 	<table border="1"> <tr> <td></td> <td>3% dd</td> <td>5% dd</td> </tr> <tr> <td></td> <td>EPID</td> <td>EBT3</td> </tr> <tr> <td>R</td> <td>-0.85</td> <td>-0.88</td> </tr> <tr> <td></td> <td>EPID</td> <td>EBT3</td> </tr> <tr> <td></td> <td>-0.78</td> <td>-0.76</td> </tr> </table>		3% dd	5% dd		EPID	EBT3	R	-0.85	-0.88		EPID	EBT3		-0.78	-0.76
	3% dd	5% dd																
	EPID	EBT3																
R	-0.85	-0.88																
	EPID	EBT3																
	-0.78	-0.76																
EAM	<ul style="list-style-type: none"> Götstedt, <i>Med. Phys.</i>, 2015 FF-IMRT and VMAT; Varian Clinac iX; Elipse; AAA; calc grid 2.5mm; Varian EPID aSi 1000 and Gafchromic™ EBT3 film 	<ul style="list-style-type: none"> Various MLC openings N=30 Pearson's correlation between dose difference and PCM value 	<table border="1"> <tr> <td></td> <td>3% dd</td> <td>5% dd</td> </tr> <tr> <td></td> <td>EPID</td> <td>EBT3</td> </tr> <tr> <td>R</td> <td>-0.94</td> <td>-0.94</td> </tr> <tr> <td></td> <td>EPID</td> <td>EBT3</td> </tr> <tr> <td></td> <td>-0.83</td> <td>-0.79</td> </tr> </table>		3% dd	5% dd		EPID	EBT3	R	-0.94	-0.94		EPID	EBT3		-0.83	-0.79
	3% dd	5% dd																
	EPID	EBT3																
R	-0.94	-0.94																
	EPID	EBT3																
	-0.83	-0.79																
C/A Götstedt, <i>Med. Phys.</i> , 2015	<ul style="list-style-type: none"> Götstedt, <i>Med. Phys.</i>, 2015 FF-IMRT and VMAT; Varian Clinac iX; Elipse; AAA; calc grid 2.5mm; Varian EPID aSi 1000 and Gafchromic™ EBT3 film 	<ul style="list-style-type: none"> Various MLC openings N=30 Pearson's correlation 	<table border="1"> <tr> <td></td> <td>3% dd</td> <td>5% dd</td> </tr> <tr> <td></td> <td>EPID</td> <td>EBT3</td> </tr> <tr> <td>R</td> <td>-0.83</td> <td>-0.84</td> </tr> <tr> <td></td> <td>EPID</td> <td>EBT3</td> </tr> <tr> <td></td> <td>-0.78</td> <td>-0.80</td> </tr> </table>		3% dd	5% dd		EPID	EBT3	R	-0.83	-0.84		EPID	EBT3		-0.78	-0.80
	3% dd	5% dd																
	EPID	EBT3																
R	-0.83	-0.84																
	EPID	EBT3																
	-0.78	-0.80																
SAS _{5mm} Crowe, <i>Australas. Phys. Eng. Sci. Med.</i> , 2014	Crowe, <i>Australas. Phys. Eng. Sci. Med.</i> , 2014 FF-IMRT, TPS Brainlab, MapCHECK2	<ul style="list-style-type: none"> Prostate N=122 beams F-value linear relationship test (p-value) 	<table border="1"> <tr> <td></td> <td>2%/2mm</td> </tr> <tr> <td>F (p-value)</td> <td>9.918 (0.002)</td> </tr> </table>		2%/2mm	F (p-value)	9.918 (0.002)											
	2%/2mm																	
F (p-value)	9.918 (0.002)																	
MAD Crowe, <i>Australas. Phys. Eng. Sci. Med.</i> , 2014			<table border="1"> <tr> <td></td> <td>2%/2mm</td> </tr> <tr> <td>F (p-value)</td> <td>0.858 (0.356)</td> </tr> </table>		2%/2mm	F (p-value)	0.858 (0.356)											
	2%/2mm																	
F (p-value)	0.858 (0.356)																	
CLS Crowe, <i>Australas. Phys. Eng. Sci. Med.</i> , 2014			<table border="1"> <tr> <td></td> <td>2%/2mm</td> </tr> <tr> <td>F (p-value)</td> <td>0.346 (0.558)</td> </tr> </table>		2%/2mm	F (p-value)	0.346 (0.558)											
	2%/2mm																	
F (p-value)	0.346 (0.558)																	
CAS Crowe, <i>Australas. Phys. Eng. Sci. Med.</i> , 2014			<table border="1"> <tr> <td></td> <td>2%/2mm</td> </tr> <tr> <td>F (p-value)</td> <td>1.818 (0.180)</td> </tr> </table>		2%/2mm	F (p-value)	1.818 (0.180)											
	2%/2mm																	
F (p-value)	1.818 (0.180)																	

(continued on next page)

Table 3 (continued)

PA Du, Med. Phys., 2014.	Du, Med. Phys., 2014. FF-IMRT/VMAT, Phillips Pinnacle, Ion Chamber Wellhoffer CC04	<ul style="list-style-type: none"> Prostate; H&N; Spine N=65 FF-IMRT + 26VMAT (prostate only) two-sided Wilcoxon rank sum test (p-value) 		Dd<3%		
			R (p-value)	-0.38 (0.0009)		
PI Du, Med. Phys., 2014.					3mm/5%	
			R (p-value)	-0.04 (0.51)		
PM Du, Med. Phys., 2014.				3mm/5%		
			R (p-value)	0.21 (0.005)		
PMU Du, Med. Phys., 2014.				3mm/5%		
			R (p-value)	0.22 (0.004)		

indicate a significant link, it was contrasted with a weak correlation. Furthermore, these results should be seen as institution-dependent and published PCM values for threshold or limits should be used carefully. The different TPS, anatomical localisation, PTQA measurement, analysis protocol or statistical correlation methods could explain these results. Therefore, with this actual state of the art, users could create their own PCM correlation for self TPS, anatomical localisation, 2D or 3D PTQA measurement, gamma criteria and threshold. So far, someone who is considering setting up a PCM as a PTQA failure identifier should quantify his own thresholds, limits and prediction capacities. Therefore, one should then consider setting up all Table 1 PCM calculations – especially MI_S, MI_a, MI_{SPORT}, M, MCS, MCS_v, MFA, CAM, EAM, SAS, MAD, CLS and PI – and then establishing correlation with their own PTQA results, with a special focus on SAS, PI, MCS and MI because of their estimated higher potential. These recommendations should allow the user to utilise a PCM as a PTQA tool and to compare his own results with the literature data.

From these previous considerations, it seems too early to affirm that a sufficient amount of data exists to prove that PCMs could replace PTQA for general cases. Nevertheless, it appears that their use can provide interesting results for clinical use [60].

4.3. Other perspectives on the use of PCMs

4.3.1. Data statistics and machine learning

Many databases relying on surveillance, epidemiology, demographic or diagnosis are now used in radiation oncology [61], and progress capacities could depend on their reliability. The objective relies on the ability to establish predictive models, and there is a strong need for high data quality to construct these bases [62]. As Mayo [63] wrote, it should be common for clinics to have the ability to rapidly gather datasets to address practice quality improvement for routine or translational research. In this perspective, one of the potential uses of PCMs could be related to the creation of dedicated databases. As the inputting information into these databases should be user-friendly and low in terms of time consumption for an easy clinical workflow implementation, it should preferably be carried out with an automatic procedure. As we are now in an era where radiation oncology data are generated daily, these databases could be employed to study the relationship between PCMs and dosimetry indices, treatment planning homogeneity, clinical outcome, mechanical component state or QA results. A data mining algorithm could also be developed to guide the analysis of these datasets and perhaps to find patterns that the human cannot detect. Even more promising, machine learning in radiotherapy [64] could be a way to develop or create new PCMs or a PCM pool that could resolve the PTQA failure problem in general cases.

4.3.2. Audit perspectives and comparison of centres

Radiation oncology [65] and, more precisely, dosimetry audits are now common processes in many countries, and they permit the evaluation of clinical practice heterogeneities between centres on themes such as TPS modelling and measurement accuracy [66,67], independent dose measurement with non-standard detectors [68] or independent peer quality control [69]. In this perspective, another potential role of PCMs could be to compare centres' propensity to deliver high or low complex plans by introducing them into these audit processes. Assessing knowledge on delivered plans with PCM score distribution would allow for a comparison of centres' practices. On the other hand, benchmarking PCM score varieties could permit one to fix limits for further clinical assay. It could also help centres to conduct an auto-evaluation of their own practices.

5. Conclusion

The aim of this work was to list and synthesise available PCMs thanks to a systematic review methodology. As the development of such variables was sensibly guided by the PTQA failure problem, a discussion about the link between available PCMs and this issue was in favour of an institution-dependent process because of the TPS, the anatomical localisation, the PTQA measurement, the analysis protocol and the statistical method dependence. So, this work plead for setting up data collection of plan PCMs in institution, and vendors as TPS manufacturers are encouraged to provide this type of tool. More specifically, SAS, PI, MCS and MI should be especially considered because of their sensitivity and their occurrence in the studies. For the use of PTQA failure identifiers, a self-PCM correlation should be made with consideration for TPS, localisation and gamma criteria of 2%/2mm. Finally, other PCM uses seem to have the potential to aid in answering this problem and to open new research fields in machine learning or audit processes.

Source of financial support/funding statement
None.

Declaration of Competing Interest

None.

References

[1] Teoh M, Clark CH, Wood K, Whitaker S, Nisbet A. Volumetric modulated arc therapy: a review of current literature and clinical use in practice. Br J Radiol 2011;84(1007):967–96.
[2] Odraszka K, et al. Clinical results of intensity-modulated radiation therapy (IMRT) for tumors of the head and neck region. Neoplasma 2005;52(2):85–94.

- [3] Cho KH, et al. Simultaneous integrated boost intensity-modulated radiotherapy in patients with high-grade gliomas. *Int J Radiat Oncol Biol Phys* 2010;78(2):390–7.
- [4] Janssen S, et al. Clinical experience of SIB-IMRT in anal cancer and selective literature review. *Radiat Oncol Lond Engl* 2014;9:199.
- [5] Khoo VS. Radiotherapeutic techniques for prostate cancer, dose escalation and brachytherapy. *Clin Oncol R Coll Radiol G B* 2005;17(7):560–71.
- [6] Daly-Schveitzer N, Julieron M, Tao YG, Moussier A, Bourhis J. Intensity-modulated radiation therapy (IMRT): toward a new standard for radiation therapy of head and neck cancer? *Eur Ann Otorhinolaryngol Head Neck Dis* 2011;128(5):241–7.
- [7] Ezzell GA, et al. IMRT commissioning: multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119: Report from AAPM Task Group 119: IMRT commissioning. *Med Phys* 2009;36(11):5359–73.
- [8] Ezzell GA, et al. Guidance document on delivery, treatment planning, and clinical implementation of IMRT: report of the IMRT subcommittee of the AAPM radiation therapy committee. *Med Phys* 2003;30(8):2089–115.
- [9] Basran PS, Woo MK. An analysis of tolerance levels in IMRT quality assurance procedures: tolerance levels in IMRT QA. *Med Phys* 2008;35(6Part1):2300–7.
- [10] McNair HA, Adams EJ, Clark CH, Miles EA, Nutting CM. Implementation of IMRT in the radiotherapy department. *Br J Radiol* 2003;76(912):850–6.
- [11] Galvin JM, et al. Implementing IMRT in clinical practice: a joint document of the American Society for Therapeutic Radiology and Oncology and the American Association of Physicists in Medicine. *Int J Radiat Oncol* 2004;58(5):1616–34.
- [12] Prescribing, Recording, and Reporting Photon-Beam Intensity-Modulated Radiation Therapy (IMRT): Contents, J. ICRU 2010;10(1):NP.3–NP.
- [13] Klein EE, et al. Task Group 142 report: quality assurance of medical accelerators: task Group 142 Report: QA of Medical Accelerators. *Med Phys* 2009;36(9Part1):4197–212.
- [14] Huq MS, et al. The report of Task Group 100 of the AAPM: application of risk analysis methods to radiation therapy quality management: TG 100 report. *Med Phys* 2016;43(7):4209–62.
- [15] Zygmanski P, Kung JH. Method of identifying dynamic multileaf collimator irradiation that is highly sensitive to a systematic MLC calibration error. *Med Phys* 2001;28(11):2220–6.
- [16] LoSasso T. IMRT delivery performance with a varian multileaf collimator. *Int J Radiat Oncol* 2008;71(1):S85–8.
- [17] Miften M, et al. Tolerance limits and methodologies for IMRT measurement-based verification QA: *Recommendations of AAPM Task Group No. 218*. *Med Phys* 2018;45(4):e53–83.
- [18] Kinsella P, McClean B, Leon-Vintro L. Demystifying failed VMAT PSQA measurements with ArcCHECK. *Phys Med* 2017;42:355–6.
- [19] Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. *Med Phys* 1998;25(5):656.
- [20] Low DA, Dempsey JF. Evaluation of the gamma dose distribution comparison method. *Med Phys* 2003;30(9):2455–64.
- [21] Jin X, Yan H, Han C, Zhou Y, Yi J, Xie C. Correlation between gamma index passing rate and clinical dosimetric difference for pre-treatment 2D and 3D volumetric modulated arc therapy dosimetric verification. *Br J Radiol* 2015;88(1047).
- [22] Heilemann G, Poppe B, Laub W. On the sensitivity of common gamma-index evaluation methods to MLC misalignments in Rapidarc quality assurance: on the sensitivity of common Gamma-index evaluation to MLC misalignments. *Med Phys* 2013;40(3).
- [23] Ford EC, Terezakis S, Souranis A, Harris K, Gay H, Mutic S. Quality control quantification (QCQ): a tool to measure the value of quality control checks in radiation oncology. *Int J Radiat Oncol* 2012;84(3):e263–9.
- [24] Nithyanantham K, et al. Influence of segment width on plan quality for volumetric modulated arc based stereotactic body radiotherapy. *Rep Pract Oncol Radiother J Gt Cancer Cent Poznan Pol Soc Radiat Oncol* 2014;19(5):287–95.
- [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] Giorgia N, Antonella F, Eugenio V, Alessandro C, Filippo A, Luca C. What is an acceptably smoothed fluence? Dosimetric and delivery considerations for dynamic sliding window IMRT. *Radiat Oncol Lond Engl* 2007;2:42.
- [27] McNiven AL, Sharpe MB, Purdie TG. A new metric for assessing IMRT modulation complexity and plan deliverability. *Med Phys* Feb. 2010;37(2):505–15.
- [28] Moher D, Liberati A, Tetzlaff J, Altman DG. for the PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339(jul211). b2535–b2535.
- [29] Anker CJ, et al. Evaluation of fluence-smoothing feature for three IMRT planning systems. *J Appl Clin Med Phys* 2010;11(2).
- [30] 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.
- [31] 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.
- [32] Li R, Xing L. An adaptive planning strategy for station parameter optimized radiation therapy (SPORT): segmentally boosted VMAT. *Med Phys* 2013;40(5):050701.
- [33] 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).
- [34] Tambasco M, Nygren I, Yorke-Slader E, Villarreal-Barajas JE. FracMod: a computational tool for assessing IMRT field modulation. *Phys Medica PM Int J Devoted Appl Phys Med Biol Off J Ital Assoc Biomed Phys AIFB* 2013;29(5):537–44.
- [35] Agnew CE, Irvine DM, McGarry CK. Correlation of phantom-based and log file patient-specific QA with complexity scores for VMAT. *J Appl Clin Med Phys* 2014;15(6):4994.
- [36] Crowe SB, et al. Treatment plan complexity metrics for predicting IMRT pre-treatment quality assurance results. *Australas Phys Eng Sci Med Sep* 2014;37(3):475–82.
- [37] 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):021716.
- [38] Park JM, Park S-Y, Kim H, Kim JH, Carlson J, Ye S-J. Modulation indices for volumetric modulated arc therapy. *Phys Med Biol* 2014;59(23):7315–40.
- [39] Park S-Y, Kim IH, Ye S-J, Carlson J, Park JM. Texture analysis on the fluence map to evaluate the degree of modulation for volumetric modulated arc therapy. *Med Phys* 2014;41(11).
- [40] Crowe SB, et al. Examination of the properties of IMRT and VMAT beams and evaluation against pre-treatment quality assurance results. *Phys Med Biol* 2015;60(6):2587–601.
- [41] Götstedt J, Karlsson Hauer A, Bäck A. Development and evaluation of aperture-based complexity metrics using film and EPID measurements of static MLC openings. *Med. Phys.* 2015;42(7):3911–21.
- [42] Park JM, Park S-Y, Kim H. Modulation index for VMAT considering both mechanical and dose calculation uncertainties. *Phys Med Biol* 2015;60(18):7101–25.
- [43] 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.
- [44] Llacer J, Solberg TD, Promberger C. Comparative behaviour of the dynamically penalized likelihood algorithm in inverse radiation therapy planning. *Phys Med Biol* 2001;46(10):2637–63.
- [45] 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.
- [46] Valdes G, Scheuermann R, Hung CY, Olszanski A, Bellerive M, Solberg TD. A mathematical framework for virtual IMRT QA using machine learning: virtual IMRT QA. *Med Phys* 2016;43(7):4323–34.
- [47] Park S-Y, Kim J, Chun M, Ahn H, Park JM. Assessment of the modulation degrees of intensity-modulated radiation therapy plans. *Radiat Oncol* 2018;13(1).
- [48] Glenn MC, et al. Treatment plan complexity does not predict IROC Houston anthropomorphic head and neck phantom performance. *Phys Med Biol* 2018;63(20).
- [49] Palaniswamy G, Scott Brame R, Yaddanapudi S, Rangaraj D, Mutic S. A statistical approach to IMRT patient-specific QA: a statistical approach to IMRT patient-specific QA. *Med Phys* 2012;39(12):7560–70.
- [50] Nelms BE, et al. Variation in external beam treatment plan quality: an inter-institutional study of planners and planning systems. *Pract Radiat Oncol* 2012;2(4):296–305.
- [51] McGarry CK, et al. The role of complexity metrics in a multi-institutional dosimetry audit of VMAT. *Br J Radiol* 2016;89(1057).
- [52] Hernandez V, Saez J, Pasler M, Jurado-Bruggeman D, Jornet N. Comparison of complexity metrics for multi-institutional evaluations of treatment plans in radiotherapy. *Phys Imaging Radiat Oncol* 2018;5:37–43.
- [53] Valentini V, Bourhis J, Hollywood D. ESTRO 2012 strategy meeting: vision for radiation oncology. *Radiother Oncol* 2012;103(1):99–102.
- [54] Malicki J, et al. Patient safety in external beam radiotherapy, results of the ACCIRAD project: recommendations for radiotherapy institutions and national authorities on assessing risks and analysing adverse error-events and near misses. *Radiother Oncol* 2018.
- [55] Malicki J. Medical physics in radiotherapy: the importance of preserving clinical responsibilities and expanding the profession's role in research, education, and quality control. *Rep Pract Oncol Radiother May* 2015;20(3):161–9.
- [56] Crowe SB, et al. Technical Note: relationships between gamma criteria and action levels: results of a multicenter audit of gamma agreement index results. *Med Phys* 2016;43(3):1501–6.
- [57] Nelms BE, et al. Evaluating IMRT and VMAT dose accuracy: practical examples of failure to detect systematic errors when applying a commonly used metric and action levels: rethinking the evaluation of dose accuracy. *Med Phys* 2013;40(11).
- [58] Woon W, Ravindran PB, Ekayanake P, S V, Lim YY, Khalid J. A study on the effect of detector resolution on gamma index passing rate for VMAT and IMRT QA. *J Appl Clin Med Phys* 2018;19(2):230–48.
- [59] McKenzie EM, Balter PA, Stingo FC, Jones J, Followill DS, Kry SF. Toward optimizing patient-specific IMRT QA techniques in the accurate detection of dosimetrically acceptable and unacceptable patient plans: comparison of patient-specific IMRT QA techniques. *Med Phys* 2014;41(12).
- [60] Kairn T, Crowe SB, Kenny J, Knight RT, Trapp JV. Predicting the likelihood of QA failure using treatment plan accuracy metrics. *J Phys Conf Ser* 2014;489.
- [61] Chen RC, Gabriel PE, Kavanagh BD, McNutt TR. How will big data impact clinical decision making and precision medicine in radiation therapy? *Int J Radiat Oncol* 2016;95(3):880–4.
- [62] Jagi R, et al. Considerations for observational research using large data sets in radiation oncology. *Int J Radiat Oncol* 2014;90(1):11–24.
- [63] Mayo CS, et al. The big data effort in radiation oncology: data mining or data farming? *Adv Radiat Oncol* 2016;1(4):260–71.
- [64] Meyer P, Noblet V, Mazzara C, Lallemand A. Survey on deep learning for radiotherapy. *Comput Biol Med Jul*. 2018;98:126–46.
- [65] International Atomic Energy Agency and Quality Assurance Team for Radiation Oncology, Eds., *Comprehensive audits of radiotherapy practices: a tool for quality improvement*. Vienna: International Atomic Energy Agency, 2007.
- [66] Budgell G, Berresford J, Trainer M, Bradshaw E, Sharpe P, Williams P. A national dosimetric audit of IMRT. *Radiother Oncol* 2011;99(2):246–52.
- [67] Clark CH, et al. A multi-institutional dosimetry audit of rotational intensity-

- modulated radiotherapy. *Radiother Oncol* 2014;113(2):272–8.
- [68] Jurado-Bruggeman D, et al. Multi-centre audit of VMAT planning and pre-treatment verification. *Radiother Oncol J Eur Soc Ther Radiol Oncol* 2017;124(2):302–10.
- [69] Lafond C, et al. DEMAT: a multi-institutional dosimetry audit of rotational and static intensity-modulated radiotherapy. *Phys Med* 2016;32(5):664–70.

Glossary

Signification of the different variables that appears in metrics PCM (Note that for a IMRT plan and/or a VMAT plan is the sum of beams PCM included in the plan weighted by Monitor Units)

- A_{Eq} : Equivalent square field or aperture
 A_k : Aperture area for the k th CP
 AAV_k : Aperture Area Variability for the k th CP (VMAT), segment (IMRT): Characterize the variation in segment area relative to the maximum aperture defined by all segments
 A_{Ik} : Aperture Irregularity calculated by considering the noncircularity of the aperture
 α : Weighting factor for the acceleration: $\frac{1}{l_k}$
 a_c : Aperture distance criteria for two opposite leaves
 a_l : Aperture distance between two opposite leaves
 CAM_k : Converted Aperture Metric for the for the k th CP
 C_{leaf} : Centre of aperture distance between opposite leaves
 C_{MLC} : Centre of the MLC axis (aligned with the beam axis)
 c : Constant in the semivariogram function $\gamma(h)$ formula
 D : Number of pairs of data points whose lag is h , for the Fractal Dimension calculation
 F : Maximum fraction of the standard deviation of the sensitive parameter considered, which represents the upper born of integration of the spectrum
 $FS(x_i)$: Fractal surface at the data point x_i
 f : Fraction of the standard deviation of the sensitive parameter considered: $f = 0.001, 0.002, \dots, 2$
 GA_k : Gantry Angle for the k th CP
 $GLCM_{i,j}$: Gray Level Co-occurrence Matrix, that indicates the intensity relationships between pairs of pixels in the fluence map
 $GLCM_i$: Mean value of the pixels in the GLCM, in the i th direction
 $\gamma(h)$: Semivariogram function used in spatial statistics that linked the Fractal Dimension (FD) to a profile of the fluence map
 $g(A_{Eq}_k)$: Conversion value of the equivalent square field, using conversion function $g(x)$, for the k th CP
 $g(d)_k$: Mean of all conversion values of the distances between MLC leaves in X and Y direction, using conversion function $g(x)$, for the k th CP
 $g(x)$: Conversion function to obtain a nonlinear relation between distance between MLC leaves, and their contribution to the complexity of the aperture
 h : Distance between two data point for the Fractal Dimension calculation
 $I_{i,j}$: Matrix of the intensity fluence map with n, m shape
 ΔI_i : Absolute difference between intensity values of adjacent-bixel in the i th direction: $\Delta I_i = |I_i - I_{i+1}|$
 K : Number of neighbouring CPs for a CP considered, which is an arbitrary value
 LSV_k : Leaf Sequence Variability for the k th CP (VMAT), segment (IMRT): Characterize the variability of a field shape per segment or CP
 $sLSV_k$: Sectoried Leaf Sequence Variability for the k th CP (VMAT), segment (IMRT): Characterize the variability of a field shape by considering the leaves that cover the organ considered in the field, per segment or CP
 λ : Weighting factor for lateral bixels, equal to 0.5 when a bixel have two lateral neighbours
 MI_i : Modulation index sensitive to intensity fluence map, which is the integration of the spectrum $Z_i(f)$
 MI_{MLC_k} : Modulation index sensitive to MLC leaf acceleration changes for the k th CP, which is the sum of all $MI_{MLC_{k,l}}$
 MI_{MLC_s} : Modulation index sensitive to MLC leaf speed changes for the k th CP, which is the sum of all $MI_{MLC_{s,l}}$
 $MI_{MLC_{a,l}}$: Modulation index sensitive to the l th MLC leaf acceleration changes, which is the integration of the spectrum $Z_{MLC_{a,k,l}}(f)$ (for one leaf only)
 $MI_{MLC_{s,l}}$: Modulation index sensitive to the l th MLC leaf speed changes, which is the

- integration of the spectrum $Z_{MLC_{s,k,l}}(f)$ (for one leaf only)
 $MLC_{a,k,l}$: MLC leaf acceleration for the k th CP (for one leaf): $MLC_{a,k,l} = \frac{|MLC_{s,k,l} - MLC_{s,k+1,l}|}{l_k}$
 $MLC_{s,k,l}$: MLC leaf speed for the k th CP (for one leaf): $MLC_{s,k,l} = \frac{|l_k - l_{k+1}|}{t_k}$
 MU : Total Monitor Units (MU) for all CPs
 MU_k : Monitor Units for the k th CP
 m : Number of bixel per column in the intensity fluence matrix (y direction)
 $\max(p_{l, \text{left bank}})$: Maximum position of the l th leaf of the left bank of the MLC
 $\max(p_{l, \text{right bank}})$: Maximum position of the l th leaf of the right bank of the MLC
 N_{CP} : Total number of control point for a VMAT plan
 N_g : Number of grey levels in the fluence map
 N_i : Total number count of intensity adjacent-bixel changes that exceed a given fraction of the intensity standard deviation of the beam
 N_{i_l} : Number count of intensity adjacent-bixel changes that exceed a given fraction of the intensity standard deviation of the beam in the i direction: $N_{i_l}(f; \Delta I_i > f\sigma_i)$
 $N_{MLC_{s,k,l}}$: Total number counts of MLC leaf speed changes for the k th CP, that exceeds a given fraction of the MLC leaf speed standard deviation (for one leaf only)
 $N_{MLC_{a,k,l}}$: Total number counts of MLC leaf acceleration changes for the k th CP, that exceeds a given fraction of the MLC leaf acceleration standard deviation (for one leaf only)
 $N_{leaf,k}$: Number of MLC leaves not positioned under the jaws for the k th CP
 $N_{leaf \text{ organ},k}$: Number of MLC leaves not positioned under the jaws, which cover the organ considered, for the k th CP
 $N_{leaf}(c > a > 0)_k$: Number of MLC leaves with an aperture distance between opposing leaves from 0 to a aperture distance criteria c for the k th CP
 $N_{leaf}(a > 0)_k$: Number of MLC open leaves for the k th CP
 $N_{leaf}(a > C_{MLC})_k$: Number of MLC open leaves that crossed the centre for the k th CP
 $N_{segment,plan}$: Total number count of segment in an particular IMRT plan (non planified with vendor recommendations)
 $N_{segment,defaultplan}$: Total number count of segment in a default IMRT plan (planified with vendor recommendations)
 n : Number of bixel per line in the intensity fluence matrix (x direction)
 P_k : Aperture perimeter the k th CP
 p_l : Position of the l th leaf of the MLC for one bank
 $p_{l, \text{leftbank}}$: Position of the l th leaf of the left bank of the MLC
 $p_{l, \text{leftbank},k}$: Position of the l th leaf of the left bank of the MLC, for the k th CP
 $p_{l,k}$: Position of the l th leaf of the MLC for one bank, for the k th CP
 $p_{l, \text{rightbank}}$: Position of the l th leaf of the right bank of the MLC
 $p_{l, \text{rightbank},k}$: Position of the l th leaf of the right bank of the MLC, for the k th CP
 p_{max} : Maximum distance between leaf positions for a MLC bank:
 $p_{os_{max}} = |\max(p_l) - \min(p_l)|$
 R_{edge} : Region of 5 mm from the MLC edge inside and outside the MLC opening
 $R_{open \text{ area}}$: Region of the MLC opening that is not taking into account by R_{edge}
 σ_i : Standard deviation of intensity values of the beam
 σ_i : Standard deviation of values of the beam, in the i th direction of the matrix
 $\sigma_{MLC_{a,k,l}}$: Standard deviation of the l th leaf acceleration of the MLC for all k th CPs
 $\sigma_{MLC_{s,k,l}}$: Standard deviation of the l th leaf speed of the MLC for all k th CP
 SLP : Slope of the $\log(\gamma(h))$ versus $\log(h)$ plot
 t_k : Time of the k th CP:

$$t_k = \begin{cases} \frac{\text{Angle gantry interval between CP}}{\text{Maximum gantry speed}} \text{ for } \Delta MU_k < \text{Maximum Dose Rate} * \text{time per CP without} \\ \text{slowing down gantry rotation} \\ \frac{\Delta MU_k}{10 * \text{Maximum Dose rate}} \text{ for } \Delta MU_k > \text{Maximum Dose Rate} * \text{time per CP without slowing} \\ \text{down gantry rotation} \end{cases}$$

 $U(A_k)$: Union area of all aperture area of a beam
 $Z_i(f)$: Total spectrum of the number of adjacent-bixel changes that exceed a given fraction of the intensity standard deviation of the beam
 $Z_{i_l}(f)$: Spectrum of the number of adjacent-bixel changes that exceed a given fraction of the intensity standard deviation of the beam, in the i direction
 $Z_{MLC_{s,k,l}}(f)$: Spectrum of the number of the MLC leaf speed changes for the k th CP, that exceeds a given fraction of the MLC leaf speed changes standard deviation for the k th CP (for one leaf only)
 $Z_{MLC_{a,k,l}}(f)$: Spectrum of the number of the MLC leaf acceleration changes for the k th CP, that exceeds a given fraction of the MLC leaf acceleration changes standard deviation for the k th CP (for one leaf only)