Physica Medica 53 (2018) 25–31

Contents lists available at ScienceDirect

Physica Medica

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

Original paper

Comparison of multi-institutional pre-treatment verification for VMAT of nasopharynx with delivery errors

E.M. Pogson[a,b,c,*], S. Arumugam[b], C.R. Hansen[d,e], M. Currie[f], S.J. Blake[a,c], N. Roberts[c,g], M. Carolan[f,g], P. Vial[a,b,c,h], T. Alharthi[a,b,c], L. Holloway[a,b,c,d,g,h], D.I. Thwaites[a]

[a] Institute of Medical Physics, School of Physics, Faculty of Science, The University of Sydney, Sydney, New South Wales, Australia
[b] Liverpool and Macarthur Cancer Therapy Centres, Liverpool, NSW, Australia
[c] Ingham Institute for Applied Medical Research, Sydney, NSW, Australia
[d] Laboratory of Radiation Physics, Odense University Hospital, Odense, Denmark
[e] Institute of Clinical Research, University of Southern Denmark, Odense, Denmark
[f] Illawarra Cancer Care Centre, Wollongong, NSW, Australia
[g] Centre for Medical Radiation Physics, University of Wollongong, Wollongong, NSW, Australia
[h] South Western Sydney Clinical School, University of New South Wales, Sydney, NSW, Australia

ARTICLE INFO

Keywords:
Gamma
3%/3 mm
Global
Simulated
Errors
MLC position
Collimator

ABSTRACT

Purpose: Measurement-based pre-treatment verification with phantoms frequently uses gamma analysis to assess acceptable delivery accuracy. This study evaluates the sensitivity of a commercial system to simulated machine errors for three different institutions' Volumetric Modulated Arc Therapy (VMAT) planning approaches.

Methods: VMAT plans were generated for ten patients at three institutions using each institution's own protocol (manually-planned at institution 1; auto-planned at institutions 2 and 3). Errors in Multi-Leaf Collimator (MLC) field size (FS), MLC shift (S), and collimator angle (C) of −5, −2, −1, 1, 2 and 5 mm or degrees were introduced.

Dose metric constraints discriminated which error magnitudes were considered unacceptable. The smallest magnitude error treatment plans deemed clinically unacceptable (typically for a 5% dose change) were delivered to the ArcCHECK for all institutions, and with a high-dose point ion chamber measurement in 2 institutions. Error detection for different gamma analysis criteria was compared.

Results: Not all deliberately introduced VMAT plan errors were detected using a typical 3D 3%/3 mm global gamma pass rate of 95%. Considering all institutions, gamma analysis was least sensitive to negative FS errors. The most sensitive was a 2%/2 mm global analysis for institution 1, whilst for institution 2 it was 3%/3 mm global analysis. The majority of errors (58/59 for institution 1, 54/60 for institution 3) were detected using ArcCHECK and ion chamber measurements combined.

Conclusions: Not all clinically unacceptable errors are detected. Combining ion chamber measurements with gamma analysis improved sensitivity and is recommended. Optimum gamma settings varied across institutions.

1. Introduction

Intensity modulated radiation therapy (IMRT) and volumetric modulated arc therapy (VMAT) both improve the conformance of dose distribution to the planning target volume (PTV) and further spare the organs at risk when compared with conventional 3D techniques. In addition, automated planning of VMAT (ap-VMAT) and IMRT has been shown to improve head and neck radiotherapy plan quality and reduce planning time relative to manual planning [1,2]. Due to the complexity of these techniques, pre-treatment dose verification is routinely utilised [3–5], whereby a measurement test is performed to verify that the intended treatment delivery for the specific patient is correct. With the advent of increasingly complex treatment deliveries, patient plan specific QA is essential [6,7]. There have been many recent studies into improving head and neck or craniospinal VMAT plan quality [8,9] and quality assurance [10,11].

There is a large variety of instrumentation and experimental methods in clinical use for pre-treatment verification. 3D anatomical dose verification has been validated from a study of 20 IMRT/VMAT planned nasopharyngeal carcinoma patients comparing to 2D phantom

* Corresponding author at: A28 – Physics, The University of Sydney, NSW 2006 Australia.
E-mail address: emp993@uowmail.edu.au (E.M. Pogson).

https://doi.org/10.1016/j.ejmp.2018.07.007
Received 9 April 2018; Received in revised form 19 July 2018; Accepted 21 July 2018

dose verification [12]. Current recommendations recognise the variation across equipment, institutions, sites and treatment planning methods and recommend the use of a pass rate calculated for an individual department as part of routine QA commissioning [13]. The accuracy requirements of linac and plan quality assurance can be found within commissioning recommendations such as TG-142 and TG-119 respectively [13,14]. Gamma criteria should specifically prioritize the property of greatest clinical importance for each treatment modality and anatomical site [3]. The advantages and disadvantages of both global and local methods are outlined in the literature [15]. Recommendations from AAPM TG 218 deem global normalisation to be more clinically relevant than local normalisation for routine IMRT QA [16]. Global normalisation has been known to mask errors at low dose points [15]. This can be considered of benefit for complex VMAT plans where the Treatment Planning System (TPS) is likely to have large uncertainties in the out-of-field low dose regions [6,17,18].

The best pre-treatment verification sensitivity occurs when using an ion chamber and planar QA method (54%) [19]. Various concerns for common QA methods have been raised throughout the literature [15,19–23] and additional factors include too lax gamma criteria, alignment accuracy issues, and differences between vendor/equipment results. The complexity of plans must be considered since complex fields can be calculated less accurately and ion chamber measurements of these plans have confirmed significant dosimetric errors [20].

This can be considered a phase III validation study according to Feygelman et al. [24], as a Phase III validation determines if the QA device is suitable for per-patient/plan dose QA by quantifying the sensitivity and specificity in detecting clinically relevant errors. There have been previous head and neck studies investigating the sensitivity of commercial detectors for linac machine errors [24–27] and TomoTherapy machine errors [28]. However none of these studies have compared the differences between departments and associated plan protocols. Liang et al. [27] investigated the sensitivity of two commercial systems, the ArcCHECK and Delta4 (ScandiDos AB, Sweden), to gantry angle, MLC position and linac output errors for head and neck cancer VMAT plans. Results from this study found the ArcCHECK to be more sensitive to gantry angle errors whilst the Delta4 was more sensitive to MLC positioning errors. Receiver Operating Characteristic (ROC) curve analysis was used in this study to avoid the choice of gamma pass rate thresholds. Fredh et al. [25] compared the sensitivities of four commercial detectors to monitor unit errors, MLC positional errors and collimator errors, concluding that different detectors were sensitive to particular errors. Arumugam et al. [29] quantified the sensitivity of the ArcCHECK detector to VMAT MLC and dose output errors. The ArcCHECK was able to detect a minimum of 3 mm MLC error and 3% output error using either a local or global gamma with 2%/2 mm tolerance. This current work rather than determining the minimum sensitivity of a detector as has already been established, aims to determine if clinically unacceptable plans (as determined by any DVH constraints outlined previously exceeding 5% [30]) are accurately detected across 3 institutions. The selected target and OAR DVH limits, typically exceeding 5% were selected in collaboration with involved radiation oncologists. A dose difference exceeding 10% for the parotid gland, in the mean dose to the parotid was used as the clinical tolerance similarly to previous studies [31–33]. Introduced errors here include MLC field size errors (opening (+) or closing (-) both MLC banks), MLC/field shift errors where MLC banks are erroneously shifted in the same direction, and collimator errors erroneously shifting the collimator angle.

One goal of this work is to compare the ideal gamma criteria for each institution. This study's overarching objective is to investigate if purposefully introduced errors of clinical significance are detected using a common commercial detector system across 3 institutions for nasopharynx radiotherapy VMAT patients, considering 4 gamma criteria with commonly used action levels and individual institution gamma action levels.

2. Materials and methods

For each error type i.e. positive MLC field sizes (FS(+)), negative MLC field sizes (FS(−)), positive MLC shifts (S(+)), negative MLC shifts (S(−)), positive collimator angle (C(+)) and negative collimator angle (C(−)) errors, the minimum error plans that were deemed clinically unacceptable from ten patients were delivered on Elekta Synergy Linacs or a Varian TrueBeam Linac using 6MV photon beams and the ArcCHECK detector within the three institutions. The cumulative dose matrix measured by the ArcCHECK was compared to the TPS-calculated dose in similar geometry using gamma analysis [5].

2.1. Baseline treatment plan and error plan generation

Volumetric Modulated Arc Therapy (VMAT) plans were generated for ten patients in Pinnacle^3 (Phillips Healthcare, Fitchburg, WI, USA) at three institutions. Each institution arranged VMAT plans from the same segmented anatomical structures of the same ten patients, shared among institutions. This dataset included one manually planned VMAT (mp-VMAT) plan (institution 1) and two ap-VMAT plans (institutions 2 and 3) for each patient. For consistency, the original Clinical Target Volume (CTV) and Organ at Risk (OAR) structures were utilized by all participating institutions for treatment planning, however it should be noted that the PTV for institution 2 was planned with their own PTV generation as per DAHANCA guidelines [34] resulting in PTVs that are contracted by 0.5 cm in all directions when compared to institution 1 and 3 which were based upon EviQ guidelines [35,36]. Differences occur as institutions were asked to plan using their own protocol. The DAHANCA contouring guideline [34] has recently been shown to improve plan consistency across multiple centres [37]. Details of the planning settings utilised in this study are given in [30]. Briefly the three institutions' nasopharynx treatment planning protocols were all planned within Pinnacle^3 version 9.10 or 9.14 with 3 dose levels of: 70 Gy to the primary tumor and gross positive lymph nodes, 59–63 Gy to high risk nodes and 54–56 Gy to low risk nodes or 68 Gy to the primary tumor and involved lymph nodes, 60 Gy to high risk nodes/regions and 50 Gy to low/elective nodal regions as per department protocol based upon ICRU83 and EviQ RU83 [35], EviQ [36] or DAHANCA guidelines [34] respectively. For specific information on included anatomical tissues within the clinical target volumes please see the respective guidelines. The dose grid resolution was set to 0.25 cm × 0.25 cm × 0.25 cm in institution 1 and 0.3 cm × 0.3 cm × 0.3 cm for the others.

Error types were intentionally introduced to the original plans utilising in-house software that has been outlined previously [26,29,38]. These error types included collimator angle (°), MLC field size (mm) and MLC shift (mm) errors of 5, 2, 1, −1, −2 and −5. Error-introduced plans were then recalculated and reviewed in MATLAB (Mathworks Incorporated, Natick USA) [30]. Plan DVH metrics were deemed unacceptable if their differences compared to the relevant baseline plan were above pre-determined tolerances, i.e. if any one or more of the following limits were exceeded; ± 5% for the PTVs D95%, brainstem D1cm^3, and spinal cord D1cm^3 or ± 10% for the parotid D_{mean}. The clinical impact of these introduced errors upon PTVs and organs at risk have been published in our previous work [30]. In this current work the smallest clinically significant error for all three error types from 3 institutions was delivered.

The delivered treatment plan details and some equipment details are outlined in Table 1. This ArcCHECK system used a 3D dose matrix in its gamma analysis, having 1386 diode detectors arranged in a helical array of 10 mm detector spacing within a Perspex cylindrical phantom [28]. The absolute dose calibration procedure recommended by the manufacturer had been followed to calibrate the device at each individual institution [39].