Reproducibility in the Field of Patient Specific IMRT QA

Elizabeth McKenzie[1], Peter Balter[1], Jimmy Jones[2], David Followill[1] , Francesco Stingo[1], Stephen Kry[1]

[1]Dept of Radiation Physics, UT MD Anderson Cancer Center, Houston, TX, [2]The Methodist Hospital, Houston, TX

Introduction

Intensity modulated radiation therapy (IMRT) has become ubiquitous in radiation clinics. The increased complexity of IMRT plans necessitates a quality assurance (QA) approach which departs from the traditional hand calculation-based verification. IMRT plans are clinically validated using direct measurement for each patient. To satisfy this need, a number of devices have been developed to measure doses from the IMRT patient plan, which is then compared to the intended dose distribution as calculated by the treatment planning system (TPS).

For the sake of convenience, several metrics have been adopted that allow for the sorting of plans as passing or failing, where a passing plan indicates that the delivered dose distribution adequately reflects the intended one (as calculated by the TPS). Two of these metrics are percent difference and percent of pixels passing the gamma criteria [1]. Percent difference is often used with point measurements, such as with an ion chamber, while the gamma analysis is used for planar measurements such as film or a diode array. The institution chooses a threshold value for these metrics to indicate whether the plan might or might not be suitable for delivery to a patient. However, the credibility of this sorting rests in part on the reproducibility in the delivery of the plan and the dose measurements. Thus, the purpose of this work is to determine the reproducibility of patient-specific IMRT QA results that one might experience clinically.

Materials

Four commercial dosimeters and one in-house designed dosimeter were selected to study their reproducibility with respect to patient specific IMRT QA. The commercial dosimeters consisted of: a Wellhofer cc04 ion chamber (CNMC, Nashville, TN), EDR2 radiographic film (Kodak Carestream, Rochester, NY), ArcCheck helical diode array (Sun Nuclear Corporation, Melbourne, FL), and a MapCheck 2D diode array (Sun Nuclear Corporation, Melbourne, FL). Additionally, the MapCheck was treated as three devices based on its analysis and delivery geometry: (1) AP field-by-field, (2) AP composite and (3) original planned gantry angles. The in-house designed dosimeter was a multiple ion chamber phantom (MIC), consisting of five separate ion chamber set in a rotational insert, allowing for multiple point measurements in 3-dimensionally independent locations. Overall, 7 dosimetry systems were considered.

Methods

Six IMRT clinical plans were chosen from thoracic, HN, GI, and GYN treatment sites, and are referred to as THOR1, HN1, GYN1, THOR2, THOR3, and GI1. These six plans were chosen to provide a variety of IMRT complexity in this reproducibility study. Each plan was delivered to a dosimetry system three times with one physical setup. The dosimeter was then removed and re-setup, and the plan was delivered again. This last step was repeated to yield a total of three deliveries under the same setup ("redelivery" reproducibility), and three deliveries under an independent setup ("composite" reproducibility containing both delivery and setup variations). This procedure is illustrated in the flow chart of figure 1. Each planar system was calculated with gamma criteria of 3%/3mm, with absolute dose (except for film which was relative).. To quantify the reproducibility, the coefficient of variation (CV) was calculated across all of the patient plans for each dosimetry system. All absolute dose measurements accounted for daily fluctuations in linac output.

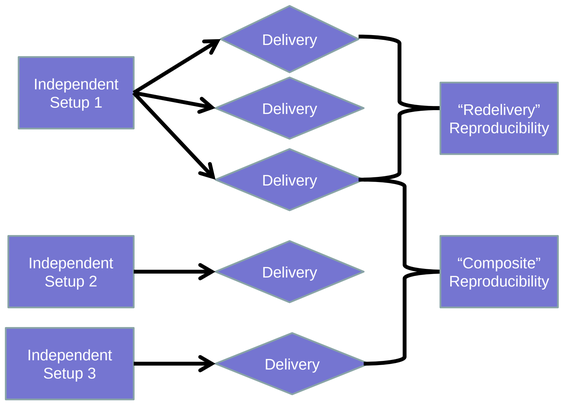
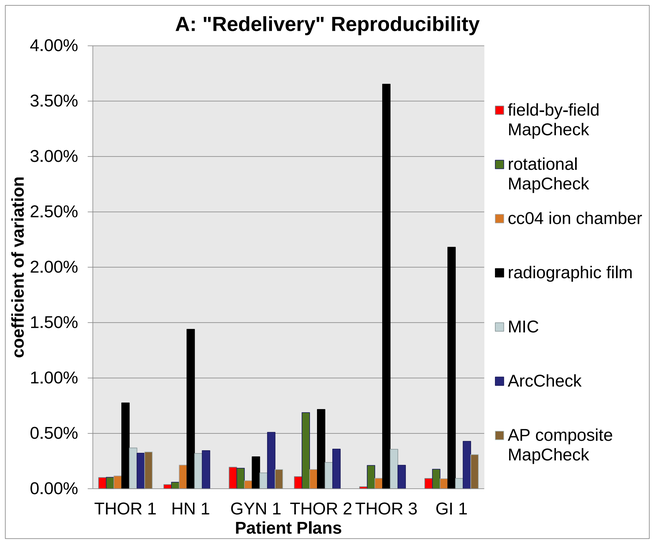


Figure 1: The workflow used to generate "redelivery" and "composite" measurements. This was repeated on each dosimeter and plan

Results

As would be expected, the variability in the "composite" measurements was higher on average than the "redelivery" measurements. This is in part because the "composite" measurements include variability from both the setup and the delivery/readout. All dosimetric systems had a CV of less than 1% for the "redelivery" reproducibility, except for film which ranged all the way to 3.7%. This may be explained by the potential sources of variability in the film readout. These include film processor conditions [2], and user selection of ROI and normalization point. Other dosimeters in this study had more immediate, less hands-on readout of the dosimetric data. For the "composite" reproducibility, film also demonstrated the most variability (average across plans of 2.0%), while the AP field-by-field MapCheck showed the least (0.15%).



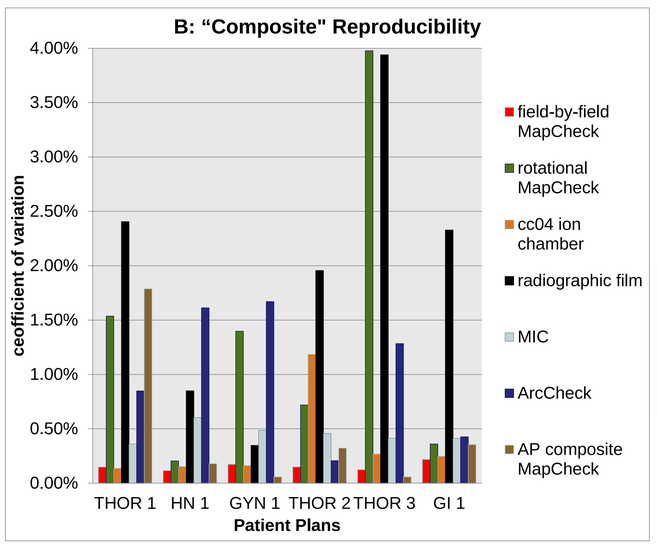


Figure 2: "Redelivery" and "composite" reproducibility expressed in terms of CV for each plan and device

Results (continued)

In order to separate the effects of the setup from the readout/delivery on reproducibility, it was assumed that these effects add in quadrature. The effect of setup was then solved according to equation 1.

$$\sqrt{\sigma_{composite\ measurements}^2 - \sigma_{redelivery\ measurements}^2} = \sigma_{setup}^2 \quad [1]$$

This analysis allows us to calculate the CV from the setup alone. It is interesting to note that most of the variation in the rotational gantry angle MapCheck comes from the setup (CV of 1.3%) compared to the delivery/readout (CV of 0.2%). Also, the film appears to have a roughly equal proportion of variation resulting from the setup alone (CV of 1.3%) and the delivery/readout (CV of 1.5%).

No plan-based statistical difference was noted after an ANOVA analysis was performed on the CV for each plan. While this suggests that reproducibility may not depend on the IMRT plan's treatment site, further measurements on a larger sample pool would be needed to confirm this.

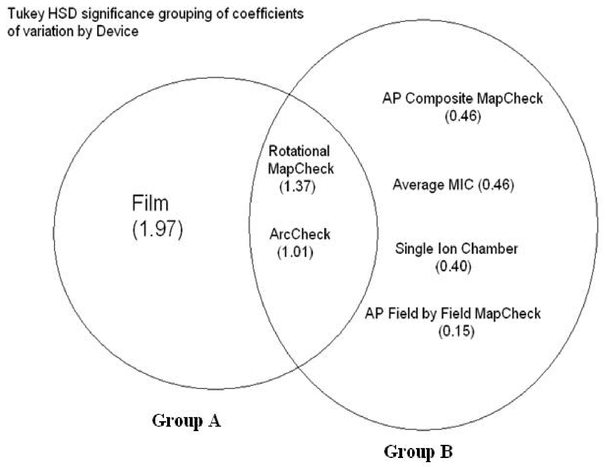


Figure 3: Venn diagram showing significant grouping of mean CV for "composite" reproducibility for each device across all six plans

Results (continued)

When the device-based results of the "composite" CV's underwent an ANOVA test, it was found that at least one group was statistically different (p-value of 0.0001). A post-hoc Tukey's HSD test was performed to assess the statistically significant grouping in the "composite" reproducibility. These results are shown in the Venn diagram of Figure 3, where film is significantly different from the AP field-by-field MapCheck, AP composite MapCheck, cc04 ion chamber, and the MIC. The MapCheck with original gantry angles and the ArcCheck were not significantly different from either group. The same analysis was conducted with the "redelivery" reproducibility, and it was found that film was the only dosimeter that was statistically different from the others.

Conclusion

A robust IMRT QA system depends in part on the reproducibility of the measured dose. This work gives the reader an idea of what kind of variability one could expect in the results if a patient specific IMRT QA measurement were retaken both with and without re-setup. Of all the dosimeters investigated here, special care should be taken with film measurements, since they are the most prone to variable results with a simple re-measurement.

With the complex gradients often found in IMRT plans, accuracy in the setup of the dosimeter could greatly influence QA results [3]. Additionally, one should be mindful of the inherent variability in the readout of the QA dosimeters. With consideration for these sources of limitations, and how they are weighted differently among dosimeters, a clinician could gain a more insightful grasp of their IMRT QA results.

References

1) Low, D.A., et al., *A technique for the quantitative evaluation of dose distributions.* Med Phys, 1998. **25**(5): p. 656-61.
2) Pai, S., et al., *TG-69: radiographic film for megavoltage beam dosimetry.* Med Phys, 2007. **34**(6): p. 2228-58.
3) Sanchez-Doblado, F., G. H. Hartmann, et al. (2007). "Uncertainty estimation in intensity-modulated radiotherapy absolute dosimetry verification." Int J Radiat Oncol Biol Phys **68**(1): 301-310.

Support

This investigation was supported by PHS grant CA10953 awarded by the NCI, DHHS.