Radiotherapy In-vivo Treatment Verification: Multi-center analysis for more than 6 million fractions delivered

Gustavo Olivera, Daniel Galmarini, Xiaohu Mo, Donald Parnell, Stephanie Key, Yu Chen, Daniel Dosoretz, Eduardo Fernandez, Constantine Mantz, Steven Finlkelstein
21st Century Oncology

Abstract

Context: Verification and safety are always a main concern in external beam radiotherapy. Even when significant time and resources are expended to guarantee quality of treatment, incidents still may happen. QA was usually performed before treatment. Verify what happen during the actual treatment may help to find undetected possible issues. With the introduction of portal dosimetry verification of behavior during treatment for every patient every fraction is clinically feasible.

Objective: Present the results of an in-vivo verification technique. Observe in-vivo deliveries and show improvements that can be achieved on the treatment consistency over time. Analyze the number of times that in-vivo verification trigger action levels and evaluate the extra workload involved.

Design, Settings and Patients: An in-vivo verification program was implemented in 84 clinics over a period of 3 years. More than 6 million daily fractions were recorded and analyze. The data cover most of the external radiotherapy manufacturers equipment (Varian, Siemens, Eleka and TomoTherapy). For every patient every fraction the calibrated portal exit dosimetry panel collected the exit beam dosimetry. An average of five-day fraction exit dosimetry panel was compared against a reference fraction using the gamma metric. A troubleshoot procedure was in place to evaluate when action level was triggered.

Results: The use of in-vivo verification allows improving the dosimetric treatment consistency as the treatment. In a period of 1 year results improved passing criteria by 6% for linacs and 7% for TomoTherapy and were maintained to those levels since then. The level of intervention was dependent on the anatomical site. Head and neck, lung and breast approximately 20% of the weekly portal triggered and action level that require verification during treatment. For pelvis 13% and less than 10% for prostate. On average between one or two average weekly exit dosimetry portals per patient need to be verified because of trigger of an action level.

Conclusions: In-vivo verification had been used to trigger lack of consistency on the exit dosimetry in more than 6 millions fractions and 84 clinics. Significant Improvements on treatment consistency are obtained and maintained by performing in vivo verification on daily basis. The number of flags detected for the more common treat anatomical site is
significant and require to be verified to clear up possible issues. Based on these numbers and difficulty in predict what patients will have flags in vivo verification should be considered for every patient every fraction. This will imply and extra work load of at least the same amount of time and sometimes even double than the time that is spend verifying pre-treatment QA (even having automatic procedures).
1- Introduction

External beam radiotherapy is a commonly used technique to treat cancer. During the last 20 years radiotherapy had rapidly evolved from three-dimensional conformal radiotherapy (3D) [1], to intensity modulated radiotherapy (IMRT)[2, 3] and many forms of image-guided radiotherapy (IGRT)[3]. These technologies aim to minimize the radiation deposited to normal tissues and in many cases allow increasing the dose delivered to the tumor. All of these techniques have comprehensive quality assurance procedures (QA) in place that had been evolving as the technology evolves[4-10]. Even with such comprehensive QA programs still incidents may happen during treatment delivery [11-23].

In IMRT the typical QA involves in phantom measurements that act as patient surrogate or by delivering the plan to a portal imager[9]. Both of them are typically performed before treatment. The IGRT portion in the majority of cases is performed by getting a portal image or a CT from the different possible on board CT capabilities available. This image capabilities aid in patient setup but also the can help during the in-vivo verification.

Verification of what happen during the time that treatment is actually delivered (in-vivo verification) may provide significant insight to evaluate the actual course of treatment. Possible issues that may be undetected with QA before treatment could be found using in-vivo verification techniques[24-26]. Countries like France[27] and UK[28, 29] have government recommendations for in-vivo dosimetry techniques. The evolution on portal dosimetry in modern machines allows to perform in vivo dosimetry in clinical routine in an efficient fashion[24, 30-38]. However, still this type of in-vivo verification may not be widely used for every patient every day, and sometimes is not even available on the treatment machines.

A lot of resources, time and equipment are invested on QA programs that in general look into what is happen before treatment. However, not always what happened during treatment is evaluated. Techniques to efficiently perform in-vivo verification increase the current workload, but also provide valuable information to evaluate and adapt (if necessary) the course of treatment.

The current manuscript presents the results from an in-vivo verification program implemented in 84 clinics where more than 6 million fractions were delivered. The in-vivo verification covers most of the machines manufactures available (Varian, Siemens, Elekta and TomoTherapy). The number of times that an intervention is flagged due to violation of an action level that triggered review during the course of treatment is analyzed for the more common anatomical sites. The intervention may correspond to possible setup, machine or anatomical changes. The number of flags decreases as function of time when in-vivo
verification was used. The verification of possible flags during treatment is a necessary condition to evaluate normal course of treatment.

2- Materials and Methods

2.1- In vivo verification technique

In order to analyze the delivery during treatment a delivery consistency procedure was implemented. Figure 1 is a workflow diagram of the implemented procedure. For each patient and each fraction a portal image of the exit dose impinging the exit detector by the transmitted treatment beam is acquired during treatment. It is worth to remark that this procedure does not imply any extra dose to the patient. The technique only collects radiation that is going through the patient.

The exit detectors for the linacs had been calibrated following manufacturer procedures in order to have their reading in Calibrated Units (CU)\[35, 39, 40\]. For the TomoTherapy machines the projection average raw data of the exit detector was used.

Patient setup for all deliveries was image guided, either with portals, CBCT or MVCT. In particular in TomoTherapy machines all patients where guided with daily MVCT. On the first implementation that we use, and for most of the results presented on this manuscript, the user selects a particular portal as reference fraction. This approach will be valid provided that the setup for that reference fraction is adequate and that the anatomical changes between the plan and actual fraction are small. After a detailed verification of the registration at the time of treatment for those two fractions, and the similarity of the anatomy between the planning CT and daily CT is checked (when CT available), one of those fractions is chosen as reference fraction if the information is consistent. If the first two fractions are not consistent, there are procedures to evaluate possible causes of the inconsistency to continue the process of selecting a reference fraction.

Once the reference fraction is selected it is compared against a computed five days average of the daily portals. To compare the consistency between the reference fraction respect to the weekly average the Gamma metric is used\[41\]. Each clinic was allowed to choose their action levels for review on between 3\% 3mm and 5\% 5mm for the linacs. In TomoTherapy was always 3\%3mm.

If the action level is violated an investigation was needed to identify and clear up the possible issues. Figure 2 is the table used during the troubleshoot procedure that includes a system review and a place to indicate possible cause of failures
determined after investigation. Also it is indicated if physics or physician action is required.

2.2- Data sets

Data had being collected for 84 clinics and for a period of 3 years. Due to the large number of clinics and the length that data had being collected not all of the data contains the same amount of detail. The number of data for linacs is described in Figure 3. Approximately 6,465,085 daily portals had being analyzed. They correspond to 1,293,085 weekly portals. For this data set it is known if they use 3% 3mm or 5% 5mm criteria and if the number of portal points passing that gamma criteria was over 95% or 90 % respectively. For approximately 4,309,000 daily portals that correspond to 861,872 weekly portals ICD9 information was available to determine anatomical location. For approximately 1,294,000 fractions that correspond to 257,858 weekly portals it is known if 3% 3mm or 5% 5mm criteria was used and the exact percentage of points of the portal imager that pass the corresponding Gamma criteria. The data set for TomoTherapy contains 42,866 daily portals.

3- Results and Discussions

Figure 4a shows the distribution of anatomical cases analyzed for linacs. Prostate, head and neck, breast, pelvis and lung encompass the main core of the portals analyzed. As mentioned earlier clinics where allowed to choose their action level for review. Some of the clinics use gamma criteria 3% 3mm and others 5% 5mm. Also, sometimes the clinics that use the 3% 3mm, as the treatment evolved if properly justified switch the criteria for the last part of the treatment to 5% 5mm. Example for this correspond to patients with significant anatomical changes and/or weight loss that may difficult the daily setup. To shed some light on the impact of using 3% 3mm or 5% 5mm as action level we analyzed the distribution of the percentage of points passing each criteria. At first glance two distinctive regions can be observed in Figure 4b. The first region (under 80 % threshold) corresponds to systematic issues such as incomplete image capture, couch bar on the exit beam, incomplete delivery, etc. The region associated to deviations link to the clinical evolution of the treatment is the region with percentage pass over 80%. For this region the distribution of points passing the criteria 3% 3mm and 5% 5mm is very similar. Also the number of cases under the 80% pass criteria is very small and decrease very rapidly. This seems to indicate that even when may be desirable to improve quality to work with the tighter criteria 3 % 3mm, still clinical relevant issues will be flag with both criteria.

Figure 5 analyze the improvements that using the in-vivo verification procedure had provided on the percentage of points passing the action level criteria from 2009 to 2012. The percentage of pass month by month is displayed for each year in each plot. It can be observed that a significant shift toward higher level of
pass (on the order of 6% on average for bins between 80 to 95%). For the subsequent years the improvements had been small and reach a plato. The use of the in vivo-verification had improve noticeable the consistency of the plan delivery and is able to keep it over time.

Figure 6 displays for each one of the anatomical sites considered the percentage of average weekly portals that pass the action level criteria. The data was analyzed separately for linacs (Figure 6a) and TomoTherapy (Figure 6b). For the linacs some cases were daily aligned by portal imagers and some other with cone beam CT. By the contrary in TomoTherapy 100 percent of the cases where CT guided and all of them with 3% 3mm tolerances. For head and neck cases 18% of the weekly in-vivo verification gamma in linacs and 22% of the weekly gamma for TomoTherapy triggered action levels for review. This means on average for head and neck cases one or may be two weeks of treatment will require a physics review to verify the adequacy of the treatment. For pelvis the percentage of weekly portal that need review is 13% for linac and TomoTherapy; for breast 26% in linacs and 16% in TomoTherapy; in lung 21% in linacs 16 % in TomoTherapy; for Prostate 11% in linacs and 4% in TomoTherapy. By no means the lower trigger percentage in TomoTherapy is associated to better treatments in TomoTherapy. For breast cases the differences are associated to the flash region. TomoTherapy does not have a flash region and linacs do. If the breast is in a different spot of the flash region for a linac may trigger action level but does not correlate to a clinical relevant issue. Also, in some few cases the retractable arms interfere in different fashions the exit beam in linacs that also is not relevant clinically. For prostate, lung and pelvis the difference between TomoTherapy and linacs is associated to interference from the couch and moveable arms at the exit beam that is not clinically relevant. Perhaps in the future threshold for flagging should become anatomically dependent.

From this figures can be inferred that on average for prostate a weekly portal every two patients will need to be reviewed. For head and neck, pelvis, breast and lung at least one and may be two weekly portals per patient will trigger action for review. Clearly different anatomical sites will need different level of controls during treatment. However, it is very important to notice that unless we measure all of them we are not able to predict what patients and when they will need review. So every patient every fraction will need to be measured and analyzed to verify at least treatment consistency. Figure 6c shows the average number of interventions needed as function of time for all of the anatomical cases analyzed. The number is around 6 % that is clearly shifted to those levels due to be prostate the majority of cases evaluated (see Figure 2). The load of review had been approximately constant over time.

Typically by just a QA before treatment many of these inconsistencies may go undetected. They will arise from possible machine, setup and/or anatomical changes. Sometimes they correspond to the normal course of treatment but sometimes no. In any case once detected and they need to be analyzed and solve if deem necessary. The role of a technique like this is to provide flags due
to lack if consistency. The clinical impact of this flags will need more ancillary process to be able to evaluate them. For instance, to be able to reconstruct the dose that was deposited on the patient is highly desirable tool to evaluate possible clinical impact of cases that triggered action levels.

In the TomoTherapy machines where 100% of the cases are CT guided we now have implemented such a tool to measure clinical impact that performs adaptive dose recalculations. The program is completely automatic and gathers the information from the same archives that the in-vivo verification does. The program extracts the daily MVCT creating a merge CT at the location where the patient was treated taking into account the image registration used for treatment. A merge image is a CT that contains the complete MVCT that was imaged with the planning kV CT filling in the rest of the image. This merge CT is used to compute the fraction dose. The dose calculator is a convolution superposition[42-44] dose calculator. To be more efficient, our implementation is GPU based. A deformable registration algorithm based on morphons[45, 46] was used to generate daily contours and analyze daily and cumulative DVHs. The deformable registration also provides deformation maps that allow mapping back the daily doses to the planning CT in order to compare planned dose respect to cumulative actual delivery. A GUI allows the user to analyze the registration, contours, daily and cumulative doses. We have implemented adaptive dose recalculation on five linacs also. As part of a future work we will also evaluate the clinical relevance of the flags provided by in vivo verification based on portals.

It is interesting to notice that the behavior per anatomical site for TomoTherapy and Linacs is very similar even when in TomoTherapy daily CT guidance is used in all cases.

From the results presented can be inferred that the implementation of a program such as the one described here, can validate consistency during treatment and also flag for verification is an action level is violated. The workload per patient (even if the process are automatic) may increase by 100 to 200 percent respect to the time that typically is expended validating pre-treatment QA. Additional work will need to be performed if an adaptive dose recalculation process is implemented to gauge clinical impact. Even when this will be an extra use of resources in a radiotherapy department, by doing this many possible issues could be detected at the more critical moment that is the treatment time.

4-Conclusions

Results from an in-vivo verification program with more than 6 millions fractions from 84 clinics were presented. The data cover most of the external radiotherapy manufacturers equipment (Varian, Siemens, Elekta and TomoTherapy).
Significant improvements on treatment consistency over time can be achieved by monitoring the behavior during treatment.

It was shown that the number of verifications that was needed due to action levels that trigger a flag was anatomy dependent. For head and neck, breast and lung approximately 20% of the weekly verification portals trigger action level. For pelvis 13% of the weekly portals required verification and prostate less than 10%. These numbers are significant to show the need to verify treatments in-vivo. Verification during treatment may detect possible issues that may go undetected otherwise. It should be observed that the level of flagging for each anatomical site was on similar levels for linacs and TomoTherapy. By one side linacs use flagging criteria of 3% 3mm and 5%5mm and the IGRT portion was either 2D portal or cone beam CT. On the other hand TomoTherapy always used 3%3mm and every fraction was CT guided.

On average between one or two weekly portals per patient trigger actions levels during the course of a patient treatment and require to be verified. The workload to do this task is considerable but we should keep in mind that may provide valuable information to QA patients. It typically will take per patient at least the same amount of time and sometimes-even double than the time that is spending verifying pre-treatment QA (even having automatic procedures).
**Figure 1:** Workflow diagram for the in-vivo verification technique used.

**Figure 2:** Table used during the troubleshooting procedure when action levels are flag.

**Figure 3:** Description of the linac data sets available for analysis.

**Figure 4:**

- **a)** Description of the case mix used on this analysis. The cases were used using ICD9 information.
- **b)** Histogram of percentage of passing points for weekly portals for the 3% 3mm and 5% 5mm criteria.

**Figure 5:** Histogram of percentage of passing points for weekly portals for each month

- **a)** 2009, **b)** 2010, **c)** 2011, **a)** 2012.

**Figure 6:** Percentege weekly portals passing in-vivo verification criteria for different anatomical sites.

- **a)** Linacs, **b)** TomoTherapy, **c)** Overall passing percentage as function of time.
Bibliography


Typical Troubleshoot Procedure

**System Review**
- Portal Vision System
  - [x] Calibration Confirmed
  - [x] Alignment Confirmed
- Treatment Plan
  - [x] Proper plan delivered
- MLC Pattern
  - [x] Proper plan delivered
- Reference Image
  - [x] Proper Image Used
- R&V System
  - [x] Proper Plan Used
- Gamma Function
  - [x] Proper Defaults Used

**Potential Causes of Failure**
- Table Location
  - Bar at Exit of Beam
  - Bar at Entrance of Beam
  - Differences on indexing
- MLC Pattern
  - Variations on Delivery
- Image
  - Incomplete Image Capture
- Treatment Delivery
  - Incomplete Delivery
- Anatomy
  - Potential Change

**Physician Action Required**
- None
- Review Setup
- Review Image
- Contact Therapist
- New Plan

**Physician Action**
- None
- New MLC QA
- Request Open BEV
- Complete Convolution
Data Sets Description

**OWAN Data**
Weekly Portals = 1,293,085
Daily Portals ≈ 6,465,000
Use criteria
3% 3mm or 5% 5mm
To meet criteria for
95% or 90%

**ICD9 Information for**
Weekly Portals = 861,872
Daily Portals ≈ 4,309,000

**Local Computers Data**
Weekly Portals = 257,828
Daily Portals ≈ 1,294,000
Know % and DTA Criteria
3% 3mm or 5% 5mm
Known Percent of Points
That meet criteria
a

Case Mix

Weekly Portals = 1,293,085
Daily Portals ~ 6,465,000

![Anatomical Site Graph]

b

Criteria Comparison Evaluation

![Graph Showing Criteria Evaluation]

Typical issues:
- Incomplete image capture
- Bar on water bath
- Incomplete delivery

Typical Course Of Treatment
Comparison Tomo and Linac

Overall Average

Clinics Average 93.9%