Original research article

Comparative analysis of image guidance in two institutions for prostate cancer patients

Tomasz Piotrowski\textsuperscript{a,c,*}, Slav Yartsev\textsuperscript{d,f,g}, George Rodrigues\textsuperscript{e,g,h}, Tomasz Bajon\textsuperscript{b}

\textsuperscript{a} Department of Medical Physics, Greater Poland Cancer Centre, Poznan, Poland
\textsuperscript{b} 2nd Radiotherapy Department, Greater Poland Cancer Centre, Poznan, Poland
\textsuperscript{c} Department of Electroradiology, University of Medical Sciences, Poznan, Poland
\textsuperscript{d} Department of Physics, London Regional Cancer Program, London Health Sciences Centre, London, ON, Canada
\textsuperscript{e} Department of Radiation Oncology, London Regional Cancer Program, London Health Sciences Centre, London, ON, Canada
\textsuperscript{f} Department of Biophysics, University of Western Ontario, London, ON, Canada
\textsuperscript{g} Department of Oncology, University of Western Ontario, London, ON, Canada
\textsuperscript{h} Department of Epidemiology and Biostatistics, University of Western Ontario, London, ON, Canada

\begin{abstract}
Aim/Background: The analysis of systematic and random errors obtained from the pooled data on inter-fraction prostate motion during radiation therapy in two institutions. Materials and methods: Data of 6085 observations for 216 prostate cancer patients treated on tomotherapy units in two institutions of position correction shifts obtained by co-registration of planning and daily CT studies were investigated. Three independent variables: patient position (supine or prone), target (prostate or prostate bed), and imaging mode (normal or coarse) were analyzed. Systematic and random errors were evaluated and used to calculate the margins for different options of referencing based on the position corrections observed with one, three, or five imaging sessions. Results: Statistical analysis showed that only the difference between normal and coarse modes of imaging was significant, which allowed to merge the supine and prone position sub-groups as well as the prostate and prostate bed patients. In the normal and coarse imaging groups, the margins calculated using systematic and random errors in the medio-lateral and cranio-caudal directions (5.5 mm and 4.5 mm, respectively) were similar, but significantly different (5.3 mm for the normal mode and 7.1 mm for the coarse mode) in the antero-posterior direction. The reference scheme based on the first three fractions (R3) was found to be the optimal one.
\end{abstract}

\begin{keywords}
Pooled database
Helical tomotherapy
Prostate cancer
Image guidance
MVCT
Margins
\end{keywords}
1. Background

Prostate cancer patients usually exhibit a considerable target motion, both between treatment fractions, primarily due to changes in the bladder/rectal filling,\(^1\)\(^2\) and during the treatment procedure itself caused by peristaltic motion and/or insufficient immobilization.\(^3\)\(^5\) With the introduction of pre-treatment imaging (e.g., megavoltage CT (MVCT) on helical tomotherapy (HT))\(^3\)\(^5\) or cone-beam CT (CBCT) on conventional linear accelerators\(^6\)), combined with the ability to verify and correct patient position with respect to the plan, the planning target volume (PTV) margin can be safely reduced\(^9\)\(^10\) allowing for more effective cancer treatment with higher prescription dose.\(^2\)\(^1\) The benefits of dose escalation have been questioned recently by Schultz and Kagan,\(^12\) but a smaller PTV margin allows for a better sparing of the sensitive organs for the same prescription dose.

Daily image guidance (IG) procedures based on MVCT or CBCT imaging are among the most effective methods of PTV reduction.\(^13\)\(^14\) However, the benefits of daily IG should be weighed against the drawbacks of increased workload for staff and in-room time and imaging radiation exposure for patients.\(^15\)\(^16\) Several groups investigated a possibility of reducing the number of imaging sessions,\(^17\)\(^25\) but relatively small patient cohorts from a single institution included in the studies limited their results.

Pooling data from different cancer centers allow increasing a database for more rigorous statistical analysis, and finally allows to obtain more precise and non-biased results. It is possible if everything is performed in exactly the same way in participating institutions. However, in most cases there are several distinctions and analysis may become quite complicated and uncertain due to a variety of possible statistical approaches with not always clearly defined application requirements. Also, there is always a question of inclusion or rejection of a patient group that intuitively is quite different from the rest. The proposal of the managing and accurate analysis of data pooling was presented in our previous paper.\(^26\)

2. Aim

The aim of this study was to analyze systematic and random errors based on pooled data from two institutions that included inter-fraction observations of prostate motion during radiation therapy. The inter-fraction patient position corrections were used to calculate the PTV margins for different conditions of the IG procedures and three options of referencing based on the position correction data from one, three, and five imaging sessions.

Conclusions: The R3 reference scheme effectively reduced systematic and random errors. Larger margins in the antero-posterior direction should be used during prostate treatment on the tomotherapy unit, as coarse imaging mode is chosen in order to reduce imaging time and dose.

© 2013 Greater Poland Cancer Centre. Published by Elsevier Urban & Partner Sp. z o.o. All rights reserved.

3. Materials and methods

An anonymized database including prospective data of 6085 megavoltage CT (MVCT) studies for 216 patients with prostate cancer was created after receiving institutional ethics approvals.

Tree sources of information to construct the final pooled database included: (1) an unpublished clinical trial database of prostate cancer palliative treatments (11 cases) at the London Regional Cancer Program (LRCP), (2) an updated version of radical treatments clinical trial database (145 cases) at LRCP,\(^13\)\(^20\)\(^21\) and (3) a clinical trial comparing dose-volume histograms for the supine and prone treatment position of the prostate cancer patients (60 cases) from the Greater Poland Cancer Center (GPCC).\(^14\)\(^27\)\(^28\)

The criteria for patient inclusion in the database were: (i) radiation treatment on a helical tomotherapy (HT) unit with daily MVCT imaging (Accuray, Madison, WI, USA), (ii) patient compliance with the preparation procedure, (iii) automatic registration of the MVCT studies to the planning kVCT studies using “Bone and Tissue Technique”, “Fine Resolution”, “Translations Only” options, (iv) availability of data on manual corrections to the automatic matching, and (v) availability of final position correction shifts applied by the radiation therapists in the lateral (x-axis), crano-caudal (y-axis), and antero-posterior (z-axis) directions.

The MVCT scanning modes were 6 mm inter-slice distance (coarse) at the LRCP and 4 mm inter-slice distance (normal) at GPCC. Coarse imaging mode was chosen at the LRCP after phantom studies on various sites\(^29\)\(^30\) and considerations of both scanning time and imaging dose reductions by 50% compared to the normal mode. No clinical assessments of different MVCT imaging options have been performed till now. All patients were asked to empty their bladder and drink 400 ml of water 1 h before the treatment and try to empty their bowels. The number of treatment fractions was 10 for palliative cases and ranging between 20 and 39 for radical cases at the LRCP, while all patients at the GPCC had 25 fractions of external beam radiotherapy followed by a brachytherapy boost.\(^31\)\(^33\) The database included the following information for each patient: (i) the number of treatment fractions; (ii) daily correction shifts in the x, y, and z directions and their manual correction components; (iii) treatment position (supine or prone); (iv) the target for irradiation (prostate or prostate bed), and (v) MVCT imaging mode (normal or coarse).

Based on the collected data, analysis of the shifts obtained by the co-registration of planning kVCT and daily pre-treatment MVCT studies with three different options for referencing was performed. At the GPCC, the patients on the first day of treatment were positioned on the external marks (tattoos) made during planning a kVCT scan; an MVCT scan...
was performed in a normal mode, these two studies were co-registered and the couch was moved to match the kVCT and MVCT studies. New external marks corresponding to the kVCT/MVCT match were marked by a fine permanent marker on the patient and these marks were used as reference for patient positioning before a MVCT scan during all consecutive treatment fractions. In this institution, correction shifts reflect changes in patient’s position with respect to the first treatment day: reference option R1. Such protocol was chosen in order to take into account systematic differences caused by: (i) different mechanical properties of the couches on planning CT and helical tomotherapy, (ii) different systems of the laser alignment on planning CT and helical tomotherapy, and (iii) image re-sampling during transfer to the tomotherapy treatment planning system. At the LRCP, patients are always positioned on external marks made during a planning kVCT study and the recorded correction shifts reflect changes in patient’s position with respect to the planning CT. In order to standardize the datasets for analysis, a modified LRCP database (LRCP-R1) was created by subtracting the correction shifts of the first treatment day from the “raw” data. Several studies have shown that 3–5 consecutive imaging sessions are enough to reduce the effective systematic error. De Boer et al. found that the initial systematic error could be reduced by 50% with 3 consecutive measurements. Bortfeld et al. proposed a “no-action level” protocol with optimum timing around 4 for 20–30 fractions. Brogi et al. determined that off-line corrections at the fourth fraction might be adequate. Beldjoudi et al. showed that residual errors decrease as the sample size for referencing grows. Therefore, two additional schemes of referencing were also evaluated to find the number of imaging fractions needed for reliable assessment of the systematic positional error: one based on average registration shifts calculated from the first three fractions (R3) and the other based on average shifts calculated from the first five fractions (R5).

The margins required for successful target coverage for these three referencing scenarios (R1, R3 and R5) were calculated using the approach of van Herk et al. The systematic error was obtained as a standard deviation of the distribution of mean errors for each patient and the random error as the average of individual patient random errors. Individual patient random errors were defined as the standard deviation (SD) of the measured errors over the course of treatment.

The statistical examination of the similarity between the groups included in the pooled database was performed using XLStat software (Addinsoft SARL, New York, NY, USA) in an MS Excel environment (Microsoft Corp., Redmond, WA, USA) according to the guidelines presented in our previous paper. The normality of the distribution was checked by the Shapiro–Wilk (SW) and the Anderson–Darling (AD) tests (when groups included less than 2000 observations) or by the Jarque–Bera (JB) and the SW tests (when groups included more than 2000 observations). If normality was not confirmed, accuracy of the fitting to the asymptotically normal distribution was checked. Possibility of the merging of groups was examined by checking the homogeneity of the variances and similarity of the averages in these groups. When normality of the distributions in the examined groups was confirmed, parametrical Fisher ($F$) and Z tests were used. In the situation when only asymptotically normal distributions were confirmed, non-parametrical Kolmogorov–Smirnov (KS) and Mann–Whitney U tests were used for examination.

Each statistical test in this study was evaluated at the significance level $\alpha = 0.05$.

4. Results

Fig. 1 shows the structure of the pooled data included in the database. Analysis of the similarity between subgroups in Fig. 1 shows that the difference between prone and supine positions for patient setup in the GPCC database was statistically insignificant, as was the case with irradiation of the prostate and the prostate bed in the LRCP database. For both situations, no significant differences between average shifts in each direction (U test for central tendency, $p > 0.05$) and similarity of the standard deviations (KS test for dispersion, $p > 0.05$) were confirmed. These results allowed us to merge the subgroups from the GPCC (prone and supine position) as well as the subgroups from the LRCP (prostate and prostate bed) and analyze them as two homogeneous groups. Therefore, our analysis was focused on the comparison between the 6 mm mode of the MVCT scans performed for 156 LRCP patients versus the 4 mm MVCT scans for 60 GPCC patients.

Fig. 2 displays the $x$ (medio-lateral), $y$ (cranio-caudal) and $z$ (anterio-posterior) correction shifts obtained for the data stratified by the imaging mode and for three referencing scenarios (R1, R3 and R5). The similarity of the central tendency between the GPCC and the LRCP group was confirmed by the U test, while the analysis of dispersion (KS test) showed statistically significant differences. The statistical description of the data in Fig. 2 is presented in Table 1.

In the analysis of the data stratified by the two different modes of imaging, the sufficient similarity leading to a possibility of merging was observed for the shifts managed by the referencing scenario (R5) in the $x$ (medio-lateral) and $y$ (cranio-caudal) directions. However, none of the referencing scenarios confirmed a possibility of merging the data for correction shifts in the $z$ (anterio-posterior) direction (Table 1).

Table 2 shows systematic ($\Sigma$) and random ($\sigma$) errors determined for different MVCT scanning modes and various referencing options, as well as the calculated van Herk margins using $2.5\Sigma + 0.7\sigma$ relation.

5. Discussion

Detailed statistical analysis based on checking the normality of the evaluated subgroups and the similarity of their central tendencies and dispersion showed the equivalence of the prone/supine subgroups in the GPCC data as well as of the tumor bed/prostate subgroups in the LRCP treatments. These results allowed us to merge the data of these subgroups and to continue with analysis of the differences between the 6 mm (LRCP) and 4 mm (GPCC) MVCT scanning options (Fig. 2 and Table 1).

The results of the U test showed that the similarity of average shifts in the $x$, $y$, and $z$ directions, while the results of the KS test for standard deviations suggest larger required margins for the 6 mm mode due to statistically significant increased standard deviations of the data obtained in this
Fig. 1 – The structure of the pooled database used in this study. GPCC – data from the Greater Poland Cancer Centre and LRCP – data from the London Regional Cancer Program.

Table 1 – Analysis with the U and K-S tests of GCPP (4 mm MVCT scanning mode) versus LRCP (6 mm scanning mode) data. Statistically significant differences are shown by bold font.

<table>
<thead>
<tr>
<th>Shift direction</th>
<th>Reference scheme</th>
<th>Average (mm)</th>
<th>U test p-value</th>
<th>SD (mm)</th>
<th>KS test p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>4 mm</td>
<td>6 mm</td>
<td></td>
<td>4 mm</td>
</tr>
<tr>
<td>x (medio-lateral)</td>
<td>R1</td>
<td>0.0</td>
<td>0.4</td>
<td>0.550</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>R3</td>
<td>-0.4</td>
<td>-0.1</td>
<td>0.512</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>R5</td>
<td>-0.3</td>
<td>0.0</td>
<td>0.491</td>
<td>1.2</td>
</tr>
<tr>
<td>y (cranio-caudal)</td>
<td>R1</td>
<td>-0.2</td>
<td>0.1</td>
<td>0.439</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>R3</td>
<td>0.3</td>
<td>0.2</td>
<td>0.444</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>R5</td>
<td>0.2</td>
<td>0.1</td>
<td>0.671</td>
<td>1.3</td>
</tr>
<tr>
<td>z (anterio-posterior)</td>
<td>R1</td>
<td>0.4</td>
<td>0.5</td>
<td>0.642</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>R3</td>
<td>0.1</td>
<td>-0.8</td>
<td>0.062</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>R5</td>
<td>0.0</td>
<td>-0.7</td>
<td>0.074</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Table 2 – Systematic, random errors and margins for 4 mm and 6 mm MVCT scanning modes.

<table>
<thead>
<tr>
<th>Shift direction</th>
<th>Reference scheme</th>
<th>Systematic (mm)</th>
<th>Random (mm)</th>
<th>Margins (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>4 mm</td>
<td>6 mm</td>
<td>4 mm</td>
</tr>
<tr>
<td>x (medio-lateral)</td>
<td>R1</td>
<td>0.9</td>
<td>2.6</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>R3</td>
<td>1.2</td>
<td>1.6</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>R5</td>
<td>1.1</td>
<td>1.6</td>
<td>3.2</td>
</tr>
<tr>
<td>y (cranio-caudal)</td>
<td>R1</td>
<td>1.2</td>
<td>1.8</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td>R3</td>
<td>1.1</td>
<td>1.2</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td>R5</td>
<td>1.1</td>
<td>1.2</td>
<td>2.2</td>
</tr>
<tr>
<td>z (anterio-posterior)</td>
<td>R1</td>
<td>1.1</td>
<td>3.3</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>R3</td>
<td>1.4</td>
<td>2.1</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>R5</td>
<td>1.4</td>
<td>2.0</td>
<td>2.5</td>
</tr>
</tbody>
</table>
Fig. 2 – Co-registration shifts obtained by matching kVCT and daily MVCT studies at the GPCC (solid points) and LRCP (open circles): (a) data for both institutions relative to the first imaging day (R1 reference), (b) data for both institutions relative to the average of the first three imaging days (R3 reference), and (c) data for both institutions relative to the average of the first five imaging days (R5 reference).
mode for the x and z directions for all referencing schemes and for the R1 reference for the y direction (Table 1). An increase of the number of imaging sessions for reference determination produced larger benefit for the 6 mm scanning mode as evidenced by a decrease of the standard deviation values in Table 1.

The values of random errors in two cancer centers shown in Table 2 are very similar as may be expected due to the same scheme of patient preparation. Maximal difference was about 1 mm. The highest differences between shifts observed in the compared groups were detected for systematic errors computed for the x and z directions on the basis of the R1 method, and were, respectively, 1.7 mm and 2.2 mm. These differences have a major influence on the disproportion observed in the sizes of margins. For example, margins in the z direction for R1 scheme should be about 6 mm larger for the 6 mm imaging group than for the group with the 4 mm method of imaging. For the R3 and R5 referencing schemes, a reduction of the differences between systematic errors for registration shifts along the x direction was observed. In the z direction case, these differences were relatively high and led to the highest difference between margins calculated for the 4 mm imaging mode (5.4 mm for the R3 and 5.3 mm for the R5 referencing methods) versus the 6 mm mode (7.4 mm for R3 and 7.1 mm for R5). Recently, Morrow et al. showed that the magnitude of set-up errors depended on the image quality. This conclusion agrees with our result on larger errors for the 6 mm mode, with a lower image quality as the image is interpolated from a larger volume. The different systematic errors in the anterio-posterior direction may also be due to different training of radiation therapists in charge and interobserver variability. A study is planned in both institutions to clarify this point.

Differences between shifts detected along the x direction may be explained by a different geometry of treatment bunkers in the two institutions. At the GPCC, the couch is on the right from the entrance and, usually, patients lie down asymmetrically on the couch, so that the radiation therapists have to move them to the left. Fig. 3 shows the deformation of the therapeutic couch due to the laying of patients always from the one side of the couch (observation at the GPCC). Before obtaining the lying position, most patients usually sit down in a similar place of the couch. The recess in the couch was confirmed by photo (Fig. 3a) and reconstruction using the MVCT images (Fig. 3c). Fig. 3b presents a transversal cross-section of the recess where the surface of the couch is about 1.5 mm lower than the non-deformed surface. The effect of the asymmetrical lying of patients can be seen in the data in Fig. 2a1 where total shifts in the x/z plane are presented. Patients at the LRCP approach the couch on their left because the entrance is on the other side of the unit and their shifts in the x direction are opposite to those of the GPCC patients (see Fig. 2a1). Using more than three fractions as the reference for position correction effectively reduces the difference between standard deviations in the GPCC and LRCP groups. Differences between these groups, e.g., registration shifts in the x and z directions, are still statistically significant (Table 1), but the calculated margins are more similar than for the R1 or R3 reference methods. We ascribed the differences in detected systematic errors in the z direction to the imaging mode because our previous study on the effect of using different immobilization devices has not shown statistically significant differences.

The use of the first method of referencing (R1) allows us to reduce errors by mechanical sources such as the deflection of the table (this phenomenon may have slightly different values for different couches). Using the average data of the first three (R3) or the first five (R5) imaging sessions as references enable to significantly reduce the differences between the standard deviations of the analyzed movements. None of the referencing methods reduces the difference between registration shifts in the z-axis measured in the two institutions. Only this difference could discriminate between the two modes of MVCT imaging (4 mm and 6 mm scans). We suggest using R3 (less workload compared to R5 with similar results)

![Image](https://example.com/image.png)

**Fig. 3** – The deformation of the therapeutic couch due to the laying of patients always from the one side of the couch. (a) Photography of the area where recess of the top of the couch was observed (in circle), (b) transversal visualization of the couch where recess about 1.5 mm was observed (arrows), and (c) the recess area reconstructed on the basis of the MVCT images. The effect was found at the Greater Poland Cancer Centre.
as a reference method in order to reduce institution specifics for the couch and bunker construction.

6. Conclusion

The 4 mm mode of the helical tomotherapy MVCT scans performed during the treatment of patients with prostate cancer was shown to produce smaller systematic error compared to the 6 mm imaging mode. The pooled data were used to calculate the planning target volume margins needed to account for inter-fraction motion for different options of referencing based on the position corrections observed with one, three, or five imaging sessions. Based on this study, the referencing scheme based on the first three fractions (R3) can be recommended with the same margins for 4 mm and 6 mm imaging mode in the x and y directions (5.5 mm and 4.5 mm, respectively) and different margins for the z direction (5.5 mm for the 4 mm imaging mode and 7.5 mm for the 6 mm imaging mode).

Conflict of interest

None declared.

Financial disclosure

This study was supported by the Greater Poland Cancer Centre (Grant No.1 dated 17 January 2012).

Acknowledgments

The authors would like to thank the staff of the GPCC Medical Physics Department and II Radiotherapy Department, in particular to Agata Jodda, Malgorzata Skorska, Adam Ryczkowski and Joanna Kazmierska for their insightful comments on this article, and radiation therapists Bartek Bak and Krzysztof Kaczmarek from the GPCC and Wendy Schoen, Diana Siegers-Krekewich, Debora Boudreau, Brittany Saldañere, June Harriman-Duke, Mona Wales, Lianne Allen, Kelly Galbraith, Abigail Owsley, Tara MacDonald, Kelly Ackerman, and Eleonore Jamieson from the LRCP for collecting data used in this work.

REFERENCES

23. Yan D, Lockman D, Brabbs D, Tyburski L, Martinez A. An off-line strategy for constructing a patient-specific planning