Effects of registration error on parametric response map analysis: a simulation study using liver CT-perfusion images

A Lausch1, NKG Jensen2, J Chen3,4, TY Lee1, M Lock3,4, E Wong1,2

1 Department of Medical Biophysics, The University of Western Ontario, London, ON, N6A 3K7, Canada
2 Department of Physics and Astronomy, The University of Western Ontario, London, ON, N6A 3K7, Canada
3 London Regional Cancer Program, London, ON, N6A 4L6, Canada
Email: alausch2@uwo.ca

Purpose: To investigate the effects of registration error (RE) on parametric response map (PRM) analysis of pre and post-radiotherapy (RT) functional images.

Methods: Arterial blood flow maps (ABF) were generated from the CT-perfusion scans of 5 patients with hepatocellular carcinoma. ABF values within each patient map were modified to produce seven new ABF maps simulating 7 distinct post-RT functional change scenarios. Ground truth PRMs were generated for each patient by comparing the simulated and original ABF maps. Each simulated ABF map was then deformed by different magnitudes of realistic respiratory motion in order to simulate RE. PRMs were generated for each of the deformed maps and then compared to the ground truth PRMs to produce estimates of RE-induced misclassification.

Main findings: The percentage of voxels misclassified as decreasing, no change, and increasing, increased with RE. For all patients, increasing RE was observed to increase the number of high post-RT ABF voxels associated with low pre-RT ABF voxels and vice versa. 3 mm of average tumour RE resulted in 18-45% tumour voxel misclassification rates.

Conclusions: RE induced misclassification posed challenges for PRM analysis in the liver where registration accuracy tends to be lower. Quantitative understanding of the sensitivity of the PRM method to registration error is required if PRMs are to be used to guide radiation therapy dose painting techniques.

1. Introduction
The ability to predict tumour response with high spatial precision during or after radiotherapy (RT) could provide for opportunities to adapt treatment to a patient’s unique tumour biology or predict their risk of recurrence. Galban et al [1] have developed a promising method for predicting tumour response by analyzing tumour functional changes on a voxel-by-voxel basis. Parametric response maps (PRMs) are generated which classify each voxel as having a significant decrease, no change, or significant increase in function. The method’s predictive utility has been demonstrated for a number of different sites and pathologies [1-4].

Ultimately, we would like to use the PRM method to develop adaptive dose painting techniques informed by repeat CT-perfusion imaging. However, PRM analysis is influenced by the ability to identify corresponding voxels in repeat scans and therefore accuracy may be reduced within imaging contexts where registration accuracy tends to be lower (e.g. abdomen). In this study we investigate...
the effects of registration error (RE) on PRM analysis by simulating known functional changes and RE on repeat CT-perfusion scans of liver tumours. PRMs are produced for different combinations of functional changes and RE and then compared to ground truth to produce estimates of RE-induced misclassification.

2. Materials and methods

2.1. Image data

The abdominal CT-perfusion scans of 5 patients with hepatocellular carcinoma were used for this study. Scans were acquired pre and post-RT using a 64 slice scanner (Discovery VCT or CT750HD, GE Healthcare, Waukesha, WI) with a spatial resolution of 0.7 mm by 0.7 mm by 5 mm and a temporal resolution of 2.8 seconds. Each image volume contained 512x512x16 voxels. Arterial blood flow (ABF) maps were generated from each 4D acquisition. A consistent image acquisition and contrast agent injection protocol was used for all patients. The imaging protocol, perfusion analysis methods, and registration technique applied to the 4D-CT scans are further described in [5].

2.2. Simulating functional changes

Seven distinct post-RT functional change scenarios (\(S_i, i = 1...7\)) were simulated per patient by modifying the pre-treatment ABF maps. Each functional change scenario is described by \(S_iP_j(x) = S_0P_j(x) + N(\mu, \sigma)\) where \(S_iP_j(x)\) is the simulated ABF map at position \(x\) for the \(i\)-th functional change scenario for patient \(j\). \(S_0P_j(x)\) is the pre-treatment ABF map at position \(x\) for patient \(j\) and \(N(\mu, \sigma)\) is a value sampled from a normal distribution with mean \(\mu = c_iT_j\) and standard deviation \(\sigma = T_j/3\). \(T_j\) is the significant change threshold for patient \(j\) and the multiplicative constants \(c_i\) vary for each scenario. The \(T_j\) were determined using the technique described in [1]. The constants \(\{c_i\} = \{0, -1.2, -1.5, -1.8, 1.2, 1.5, 1.8\}\) were selected to produce a spectrum of different PRM classification outcomes which predict no change (\(c_i = 0\)) or small, medium and large tumour response (\(c_i < 0\)) and progression (\(c_i > 0\)). For simplicity, tumour size was left unchanged in the simulated post-RT ABF maps. Each \(S_iP_j\) was blurred using a 3x3x1 averaging filter to reduce pixelation. An example of the ABF functional change simulations is shown in Figure 1A.

Functional changes were simulated under the premise that high-ABF tumour regions are more likely to decrease in function after successful RT and low ABF-regions are more likely to increase in function after unsuccessful RT. Consequently for response and progression scenarios, tumour voxel values within the top 70% and bottom 70% of all tumour ABF values were modified using the non-zero \(c_i\) while the remaining 30% of voxels were modified using \(c_i = 0\). At the conditional boundary, \(c_i\) was scaled linearly between zero and the non-zero values to eliminate ABF value discontinuity.

2.3. Simulating image registration error

In this study, simulated deformable motion was used to emulate the residual image misalignment caused by RE. Each simulated ABF map \(S_iP_j\) was deformed to produce twenty-one copies \((S_iP_j M_k, k = 0...20)\) associated with different magnitudes of inter-scan RE. Abdominal structures move in the anterior and inferior directions when a patient lies supine [6]. To simulate this motion, the diaphragm was modeled as a half sinusoid with varying amplitude similar to [7]. \(S_iP_j M_k\) was generated by displacing \(S_iP_j\) voxels according to,

\[
\Delta SI(x, k) = -\Delta SI_{max} \sin(\pi k/k_{max}) \sin(\pi x/x_{max})
\]
\[
\Delta AP(x, k) = \Delta AP_{max} \sin(\pi k/k_{max}) \sin(\pi x/x_{max})
\]

where \(\Delta SI_{max} = 10\) mm and \(\Delta AP_{max} = 5\) mm were the maximum SI and AP displacements respectively. \(x_{max}\) is the LR extent of the image volume, and \(k_{max} = 20\). Finally, a small rotation
about the SI axis was added and was specified by \( \theta(k) = \theta_{\text{max}} |\sin(\pi k/k_{\text{max}})| \) where \( \theta_{\text{max}} = 1.5^\circ \).

Figure 1B and 1C illustrates an example of the influence of simulated RE on the ABF maps.

2.4. Parametric response maps

PRMs are generated by comparing two functional imaging scans to one another. A voxel of tissue is classified as significantly increasing or decreasing if its change in function (\( \Delta F \)) is such that \( \Delta F > +T \) or \( \Delta F < -T \) respectively, where \( T \) is a threshold determined by investigating the change in function within similar healthy tissue [1]. Here, the threshold for each patient (\( \pm T_f \)) was determined by comparing their pre and post-RT ABF maps within the healthy liver. The maps were first non-rigidly registered to one another by applying a transformation derived from registering the average of the 4DCT scans used to create each map. Elastic, multi-resolution registration was performed using the Flexible Algorithms for Image Registration Toolkit [8]. These thresholds were used to inform functional change simulations since functional changes derived from a direct comparison of the pre and post-treatment ABF maps would be confounded by RE. The ground truth and RE influenced PRMs were generated by comparing \( S_0P_jM_k \) to \( S_0P_j \) and \( S_0P_jM_k \) to \( S_0P_j \) (\( k \neq 0 \)) respectively.

3. Results and discussion

Similar to [1], scatter plots of \( S_0P_jM_k \) values versus corresponding \( S_0P_j \) values were created to visualize changes in tumour ABF values with respect to \( \pm T_f \). Figure 2 illustrates representative plots for patient 3, where \( k = 0 \) (no RE) and \( k = 20 \) (~7.5mm average tumour RE). Average tumour RE refers to the average Euclidean distance between corresponding tumour voxels within compared scans. The plots for \( k = 0 \) represent the ground truth functional changes.

The percentage of voxels classified as increasing/decreasing deviated away from the ground truth as RE increased. For all patients, increasing RE was observed to increase the number of high post-RT ABF voxels associated with low pre-RT ABF voxels (and vice versa) within compared maps. This phenomenon produced a characteristic artifact within the scatter plots which appeared as a spreading out of the points contained by a triangular shaped envelope. In practice, this phenomenon may be difficult to distinguish from scenarios where parts of the tumour with low functional values actually increase and parts with high functional values actually decrease.

The true and RE influenced PRMs were compared to compute the percentage of tumour voxels that were misclassified due to varying magnitudes of RE. The average percent misclassification among the 5 patients was computed for all combinations of functional change scenarios and simulated RE and was then plotted against the average tumour voxel RE (figure 3).
For all scenarios, misclassification steeply increased and then gradually plateaued. We hypothesize that the severity of the increase is influenced by the interplay between the magnitude of RE and the size and heterogeneity of the features within the compared ABF maps. For example, if features in the compared ABF maps are very small and heterogeneous (particularly in the direction of maximum RE), then small amounts of RE can lead to complete misalignment and result in very high rates of misclassification. As the RE exceeds the feature size, the misclassification rate begins to plateau since feature misalignment has been maximized. Consequently the exact relationship between misclassification and RE will differ for every tumour depending on feature size and heterogeneity within the interrogated functional maps. This helps to explain the lower misclassification rates observed for the response and progression scenarios since these scenarios were defined in a manner 

**Figure 2.** PRM scatter plots, \( S_p M_k \) versus \( S_0 P_3 \) tumour ABF values (normalized). \( k = 0, 8, \) and 20 correspond to 0 mm, 4.3 mm, and 7.5 mm of average tumour RE respectively. The hatched lines indicate the significant change thresholds. The percent of tumour voxels classified as significantly increasing/decreasing is also indicated.

**Figure 3.** Average percent misclassification of tumour voxels due to registration error.
which increased tumour ABF homogeneity. Increased homogeneity decreases classification sensitivity to RE since features are effectively increased in size.

4. Conclusion
In this study we have simulated known functional changes and RE to investigate the effects of RE on parametric response map analysis using liver tumour CT-perfusion images. 3 mm of average tumour RE was observed to result in 18-45% voxel misclassification rates. This presents a major challenge towards the development of PRM-guided adaptive dose painting techniques since RE-induced misclassification could lead to incorrect radiation targeting. However, this challenge is not unique to the PRM response prediction method. All approaches involving a voxel-to-voxel comparison of registered images could be highly sensitive to RE depending on the interplay between RE magnitude/direction and the size, complexity, and orientation of features within the analyzed images. Since this interplay is unique to every image set and registration, methods need to be developed which can characterize and potentially mitigate RE-induced uncertainty in the analysis of each image voxel. The analysis uncertainty can then be visualized and taken into consideration during treatment planning. Without this information, voxel-based analyses like the PRM should not be used to guide locally adaptive dose painting unless it has been carefully verified that the maximum registration error within the images is much smaller than the image features.

References
[1] A Galban C J et al. 2009 Nat. Med. 15(5) 572-76
[2] A Galban C J et al. 2009 Transl. Oncol. 2(3) 184-90
[3] M Hoff et al. 2012 Bone 51(1) 78-84
[4] Harris R J, Cloughesy T F, Pope W B, Nghiemphu PL, Lai A, Zaw T, Czernin J, Phelps ME, Chen W, Ellingson BM. 2012 Neuro-Oncology 14(8) 1079-89
[5] Jensen N K G, Lock M, Fisher B, Kozak R, Chen X, Chen J, Wong E and Lee TY. 2013 Acad. Radiol. 20(4) 414-22
[6] Brandner E D, Wu A, Chen H, Heron D, Kalnicki S, Komanduri K, Gerszten K, Burton S, Ahmed I, Shou Z. 2006 Int. J. Radiat. Oncol. Biol. Phys. 65(2) 554-60
[7] Lausch A, Ebrahimi M, Martel A 2011 Image registration for abdominal dynamic contrast-enhanced magnetic resonance images. Proc. 8th IEEE Int. Symposium on Biomedical Imaging: From Nano to Macro (Chicago, USA, March 30-April 2 2011). pp 561-65
[8] Modersitzki J 2009 FAIR: Flexible Algorithms for Image Registration. Philadelphia, USA: Society for Industrial and Applied Mathematics.