

Comparison of rigid and deformable image registration accuracy of the liver during long-term transition after proton beam therapy

Objective: To identify the previously irradiated region is important in re-irradiation. We investigated the long-term transition of geometrical accuracy of a commercially available rigid and deformable image registration (RIR, DIR) software product, MIM Maestro, in livers of patients who underwent proton beam therapy (PBT).

Methods: Image registration of pre-treatment and post-treatment CT was performed in 15 liver tumor patients who received PBT and follow up post-treatment CT examination was performed over time (2-7 times each, a total of 54 times). We performed RIR and DIR of pre-treatment CT and compared the post-treatment CT by calculating the similarity of whole liver (Dice similarity coefficient: DSC) and dislocation of metallic markers implanted close to the tumor (fiducial registration error).

Results: In the plain CT, the DSC was bigger with RIR in 1 case and DIR in 53 cases (0.83 ± 0.07 vs. 0.92 ± 0.04), then consistently bigger with DIR in 14 patients. The fiducial registration error was smaller with RIR in 10 cases, DIR in 43 cases, and equal error in 1 case (11.8 ± 6.9 vs. 7.6 ± 9.7 mm), then consistently smaller with DIR in 10 patients. In the contrast-enhanced CT, the DSC was consistently bigger with DIR in all 15 patients, 54 cases in all (0.84 ± 0.06 vs. 0.92 ± 0.03). The fiducial registration error was smaller with RIR in 7 cases and DIR in 47 cases (11.0 ± 6.3 vs. 6.3 ± 7.3 mm), then consistently smaller with DIR in 12 patients.

Conclusion: The DIR performance of MIM is therefore superior to that of RIR and its advantage is independent the term after PBT.

KEYWORDS: deformable image registration ▪ rigid image registration ▪ geometrical accuracy ▪ Liver ▪ proton beam therapy ▪ dice similarity coefficient ▪ fiducial registration error

Introduction

Although irradiation of liver cancers is a well-established treatment modality since proton beam therapy (PBT) was developed [1-4], patients often develop metastatic lesions in the liver several months or years later and additional treatment is necessary. It is vital to confirm previously PBT-irradiated regions to avoid excess irradiation to normal tissue as this could cause severe side effects. Because PBT has excellent dose concentration than photon radiotherapy, dramatic inclined deformation of the liver should be occurred. Moreover, several times of re-irradiation can be expected because irradiation dose to the normal liver can be limited to a lesser extent [5,6]. Therefore, many opportunities to be bothered in identifying the previously irradiated region correctly.

Several Deformable Image Registration (DIR) algorithms exist, and DIR can be broadly classified into two 2 categories: (1) intensity-based methods, which use a variety of image

intensity metrics such as gray scale or (2) feature-based methods which use specific image features such as contours [7]. Transformation models include optical flow-based equations [8], “demons” equations [9], b-splines [10] and thin plate splines [11]. In most registration algorithms, the balance between image similarity and accurate matching of local features on one hand and deformation smoothness on the other hand is crucial to accurately measure the deformation [12].

In recent years, advanced DIR software packages have been developed for research purposes, and some of them are also available for clinical use. MIM Maestro (MIM Software Inc., USA) (MIM) is the most widely used, globally available DIR software that can assist in radiotherapy planning [13,14]. In MIM, the DIR algorithm is an intensity-based, freeform cubic spline interpolation with essentially unlimited degrees of freedom [12,15,16]. Intensity-based algorithm deforms following change of each

Nobuyoshi Fukumitsu*,
Toshiyuki Terunuma,
Toshiyuki Okumura,
Haruko Numajiri, Keiko
Murofushi, Kayoko
Ohnishi, Masashi
Mizumoto, Teruhito
Aihara, Hitoshi Ishikawa,
Koji Tsuboi & Hideyuki
Sakurai

Proton Medical Research Center,
University of Tsukuba, Tsukuba, Japan

*Author for correspondence:

fukumitsu@pmrc.tsukuba.ac.jp

pixel value and free form interpolation enables high flexibility of deformation. Thus, MIM is expected to be useful for comparing the same modality and can follow a big change, while has a disadvantage of excessive deformation, although the respective algorithms are only partially published and detailed information about them is still undisclosed.

In our previous study, we revealed the fiducial registration error of DIR is 9.3 ± 9.9 mm at one point analysis between 4 and 14 months after PBT and approximately 3 mm smaller than rigid image registration (RIR) [14]. The size and shape of the liver changes during long-term transition after PBT. Thus, both DIR and RIR accuracy is expected to be reduced gradually. Although DIR has a function to complement the shortcoming of RIR, occasionally cause excessive deformation by the artefact or effusion. Thus, it may be possible that RIR is rather advantageous in case of tiny deformation by avoiding the risk of excessive deformation. Also, with too much deformation, it is worried that DIR would cause unnatural deformation and reliability of DIR and RIR may be reversed beyond the limit of DIR adaptation. Next our clinical concern is proper use of RIR and DIR. That is when or how much level of deformation reverses the superiority of image registration accuracy between RIR and DIR. Thus, we conducted an examination of the RIR and DIR geometrical accuracy function of MIM in PBT-irradiated livers that were followed long-term transition after PBT.

Materials and Methods

We retrospectively reviewed patients who had received PBT at our institute. All the study procedures involving human participants were conducted in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. All the treatments were discussed at in hospital conferences and informed consent was obtained from all the individual participants included in the study. The study received institutional review board approval (H28-102). We selected those who had a metallic material (e.g. a fiducial marker or surgical clip; hereafter called "metallic marker") implanted close to the tumor in the liver. We examined 15 consecutive patients between 2009 and 2014 (age range:

54-80 years old; 11 men and 4 women). The most common disease was hepatocellular carcinoma (10 total), liver metastasis (4 total) and intrahepatic bile duct carcinoma (1 total). Metallic markers for previous PBT were present in 9 patients and surgical clips in 6. In our institute, abdominal CT for diagnosis is usually not taken after metallic marker implantation, so these 9 patients had come to our hospital to receive PBT for new lesions in the liver. Total tumor diameter was 20-65 mm (median: 35 mm). The distance between the tumor and metallic marker was 5-33 mm (median: 12.0 mm). Total irradiation dose was 50-74 GyE in 20-37 fractions.

Plain and contrast-enhanced CT with a breath-holding technique was taken before and after treatment. Post-treatment CT scans were taken over time, such as the duration between PBT and first post-treatment CT was 2-7 months and last post-treatment CT was 7-24 months. The number of post-treatment CT examinations was 2-7 times (median: 3) with a total of 54 times. CT with a matrix resolution of 512×512 pixels and a slice thickness of 5 mm was used. We used both plain and contrast-enhanced CT in our image analysis. For the patients who had dynamic contrast-enhanced CT, we used portal venous phase CT. We performed RIR and then DIR of the pre-treatment CT images to the post-treatment CT images.

Contour of the whole liver and dislocation of the metallic marker was examined using MIM (version 6.5.2). Dice similarity coefficient (DSC) was used for comparison of the whole liver. The DSC is defined as

$$DSC = \frac{2|A \cap B|}{|A| + |B|}$$

The values of the DSC range from 0 to 1 and the identical to 1 if the A: pretreatment liver volume and B: posttreatment liver volumes are equal with complete intersection [17]. Moreover, liver volume was measured and posttreatment large liver volume change was defined as greater than 20% difference compared to the pretreatment CT. The fiducial registration error was assessed by the examining the discrepancies in the location of the metallic marker from its posttreatment position to its pre-treatment position after RIR or DIR. We used the highest density point in the marker as the position of the metallic marker.

The value was presented as mean ± standard deviation. The fiducial registration error of MIM was compared between pre-treatment and post-treatment CTs using paired t-tests. Statistical analysis was performed using Statcel 4. Probability values below 0.05 were considered significant.

Results

In the 54 cases of plain CT, the DSC range was 0.65-0.94 (0.83 ± 0.07) in RIR and 0.81-0.99 (0.92 ± 0.04) in DIR (P=8.6 × 10⁻¹⁷) and the biggest DSC was RIR in 1 trial and DIR in 53 trials. The fiducial registration error range was 2.2-27.6 mm (11.8 ± 6.9 mm) in RIR and 0.2-40 mm (7.6 ± 9.7 mm) in DIR (P=4.0 × 10⁻⁴) and the smallest fiducial registration error was RIR in 10 trials, DIR in 43 trials, and 1 trial showed no difference between the two methods. In the cases of contrast-enhanced CT, the DSC range was 0.66-0.94 (0.84 ± 0.06) in RIR and 0.85-0.97 (0.92 ± 0.03) in DIR (P=3.1 × 10⁻¹⁸) and the biggest DSC was DIR in all 54 trials. The fiducial registration error range was 1.1-26.9 mm (11.0 ± 6.3 mm) in RIR and 0.1-32.6 mm (6.3 ± 7.3 mm) in DIR (P=9.3 × 10⁻⁷) and the smallest fiducial registration error was RIR in 7 trials and DIR in 47 trials (FIGURE 1).

TABLE 1 shows the change in the accuracy superiority of image registration during follow up. In the plain CT, the DSC was consistently bigger in DIR in 14 patients, and inconsistent in 1 patient, and fiducial registration error was consistently smaller in RIR in 1 patient, in DIR in 10 patients, and inconsistent in 4 patients. In the contrast-enhanced CT, the DSC was consistently bigger in DIR in all 15 patients and the fiducial registration error was consistently smaller in RIR in 1 patient, in DIR in 12 patients and inconsistent in 2 patients.

Massive fluid collection was observed in 12 cases. The liver volume was changed as 66-123 (94 ± 11)% to the pre-treatment CT. Large liver volume change was observed in 6 cases. Fluid collection or large liver volume change were observed in 10 of 15 cases whose DIR fiducial registration error beyond 10 mm in the plain CT, 7 of 9 cases in the contrast-enhanced CT, all 5 cases whose DIR fiducial registration error consistently bigger than RIR or inconsistent in the plain CT and 2 of 3 cases in the contrast-enhanced CT (FIGURE 2).

FIGURE 3 shows cases whose fiducial registration error was consistently smaller with DIR. The metallic marker was not found with RIR, but could be found in DIR 5 and

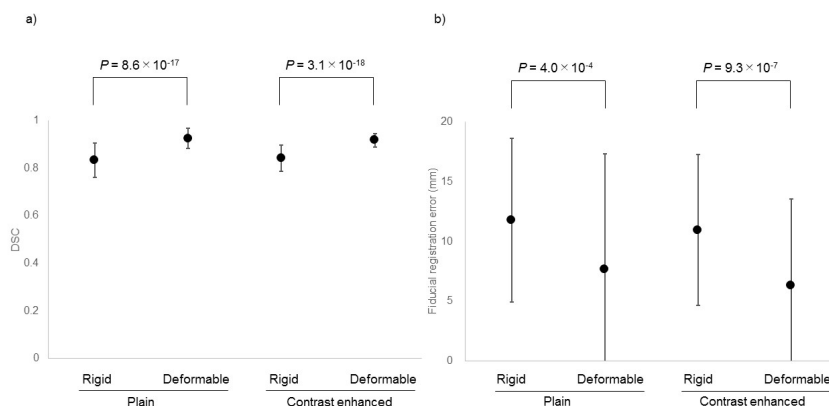


Figure 1. DSC and fiducial registration error. (a) DSC (b) fiducial registration error. Data represents mean value and standard deviations.

Table 1. Comparison between RIR and DIR during follow up.		
	DSC RIR bigger/DIR bigger/ Inconsistent	Fiducial registration error RIR smaller/DIR smaller/ Inconsistent
Plain	0/14/1	1/10/4
Contrast enhanced	0/15/0	1/12/2

Abbreviations: RIR: rigid image registration; DIR: deformable image registration; DSC: dice similarity coefficient

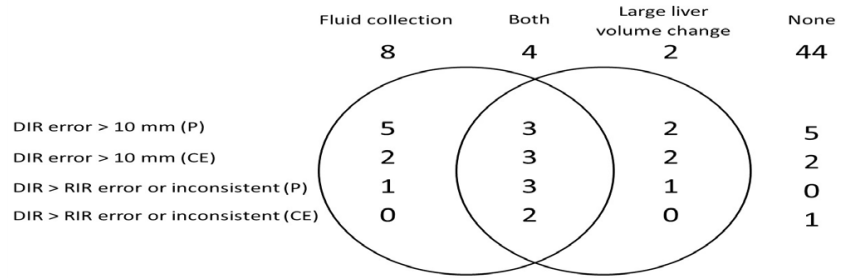


Figure 2. Correlation with unstable DIR fiducial registration error (fluid collection and liver volume). DIR: deformable image registration; RIR: rigid image registration; P: plain; CE: contrast-enhanced

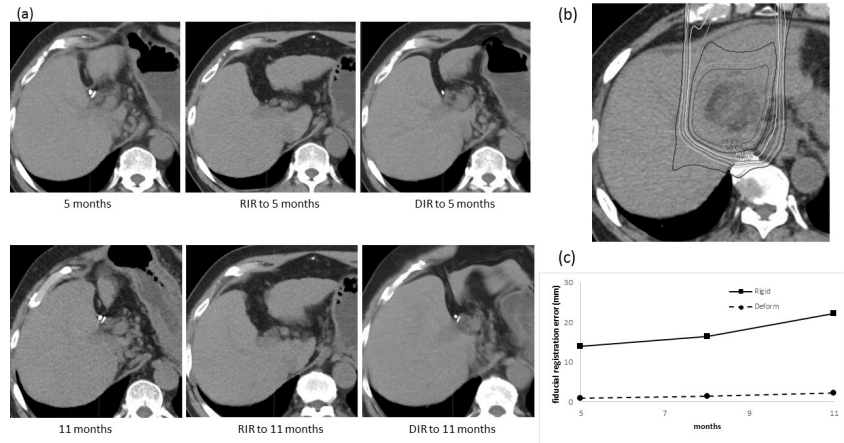


Figure 3. 68 year old man with HCC in S4. (a) Upper: 5 months later (Left: post-treatment CT; Middle: RIR of pretreatment CT to post-treatment CT; Right: DIR of pre-treatment CT to post-treatment CT). Lower: 11 months later. (b) Dose distribution of PBT. Isodose lines represent 95% to 10% from inside to outside. (c) The graph represents the transition of RIR (straight line) and DIR (dotted line).

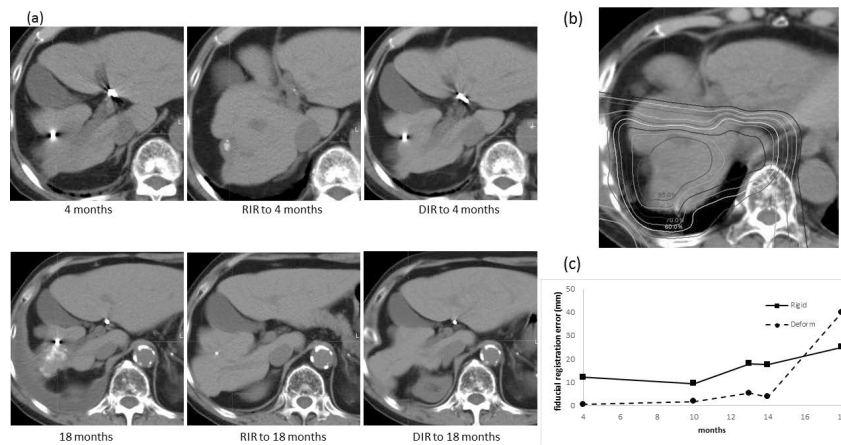


Figure 4. 74 year old man with HCC in S5/8. (a) Upper: 4 months later. Lower: 18 months later. (b) Dose distribution of PBT. (c) Transition of RIR and DIR.

11 months later. FIGURE 4 shows cases whose DIR fiducial registration error showed inconsistencies. The metallic marker was barely visible with RIR but quite clearly visible with DIR at 4 months. However, after 18 months, the marker was not definitively found with DIR but was barely visible with RIR. Massive pleural

effusion was found in CT 18 months later and excess deformation of the thorax was found with DIR.

Discussion

We utilized a one-point, CT-based analysis in 24 patients with measurement times between 4

and 14 months. The fiducial registration error was 1.6–42 mm (12.5 ± 9.4 mm) with RIR; 0.4–32.9 mm (9.3 ± 9.9 mm) with DIR in the plain CT and 3.1–27.2 mm (12.2 ± 7.2 mm) with RIR; 1.0–24.9 mm (7.4 ± 7.7 mm) with DIR in the contrast-enhanced CT [14]. We investigated when or how much level of deformation reverses the superiority of image registration accuracy between RIR and DIR in this study. The fiducial registration error in the range of a 2 years follow-up was not so different from that of our previous one-point analysis in total. The fiducial registration error was consistently smaller with DIR in 10 out of 15 patients in the plain CT and 12 patients in the contrast-enhanced CT. Moreover, the value of the fiducial registration error of RIR - DIR went upward in 6 patients, downward in 8 patients and flat in 1 patient in the plain CT while the contrast-enhanced CT showed higher error in 8 patients and lower error in 7 patients, which showed no clear trend. In regard to the whole liver registration, the DSC was consistently bigger in DIR in most of the patients (14 out of 15 patients in the plain and all 15 patients in the contrast enhanced CT). Therefore, we consider that geometrical accuracy of the liver by using DIR is superior to that of RIR usage and this advantage is not necessary affected by the term after PBT. To put this in perspective, however, it should be noted that the patients having a massive fluid collection or large liver volume change during follow ups do not necessarily show consistently smaller fiducial registration errors of DIR. This indicates that the geometrical accuracy of DIR sometimes may become unstable by excessive anatomical changes.

MIM is relatively new and, therefore, there are very few reports about the utility of DIR. Some of those reports, however, contain some very useful information. Nie and colleagues reported that MIM has an advantage in low-contrast small regions and is heavily influenced by noise [15]. Kirby and colleagues reported that, although it is difficult to indicate which algorithm is the best, MIM either adjusts to the image noise or it can produce a large error [12]. Singhrao and colleagues reported that MIM produces beautiful image similarity, but may produce nonphysical deformation fields [16]. To summarize, MIM has the advantage for a small field registration. It has higher flexibility but is easily affected by noise and artifacting

which may cause unreasonable deformation. To the best of our knowledge, this is the first study which examined how the accuracy of DIR changes during follow up. Thus, we do not have any comparable previous studies to ours. However, it is common knowledge within our field that unreasonable deformation is sometimes found. Moreover, the fact that the fiducial registration error in the contrast-enhanced CT was smaller than plain CT means that the intensity-based algorithm of MIM [12] satisfactorily helps to improve DIR of the liver tissue which has a relatively homogeneous Hounsfield unit ranking.

It may have criticism about technique measuring registration accuracy using a metallic marker. The whole liver registration accuracy using the DSC was additionally investigated to respond to the criticism and to increase the credibility of the data. The DSC can evaluate the image registration accuracy of the whole liver and fiducial registration error is characterized by being able to evaluate position matching of remarkable area of deformation near the tumor with accuracy of millimeter. The metallic artifact was so small to affect the accuracy of the image registration of the large volume of the liver. The number of implanted metallic markers is usually 1 or 2 in daily clinic. Thus, we selected only the patients whose metallic materials were close to the tumors with a range of 5–33.7 (median: 12.0) mm in this study as already noted in our previous study [14]. This analysis reflects the conditions of the daily clinical setting, and we consider that analysis using whole liver and a tiny metallic marker is clinically allowable. It may have criticism about large error of DIR (7.6 ± 9.7 mm in the plain CT and 6.3 ± 7.3 mm in the contrast-enhanced CT) is not clinically tolerable. We understand completely correct method to identify the previous irradiation region in several months or years is not existed. However, it is required for decision and delineation of the previously irradiated region at reirradiation. We usually refer to both of RIR and DIR in the clinical setting. The important thing is to clarify which image registration is more reliable. We consider our data that DIR is more reliable than RIR and the trend is not reversed during long-term transition are clinically important.

MIM has an extensive feature set for radiotherapy planning assistance that has

rapidly expanded. For example, MIM has a "Reg Refine" tool which allows users to voluntarily deform the image, but we did not use this tool because the results are based solely on operator technique and data objectivity can be easily lost. There is no doubt that future versions will have enhanced versatility and each study will have to take new tools and improved DIR accuracy into consideration when planning a study.

Conclusion

For image registration in the liver, the DIR

performance of MIM after PBT is superior to RIR, and the trend is independent the term after PBT. However, large anatomical changes sometimes reduce the DIR performance of MIM.

Acknowledgement

The authors would like to thank Dr. Bryan Mathis of the Tsukuba University Medical English Communications Center for proofreading our manuscript. This study is partially supported by JSPS KAKENHI Grant Number 17H04256.

REFERENCES

- Bush DA, Kayali Z, Grove R *et al.* The safety and efficacy of high-dose proton beam radiotherapy for hepatocellular carcinoma: A phase 2 prospective trial. *Cancer*. 117, 3053-3059 (2011).
- Fukumitsu N, Sugahara S, Nakayama H *et al.* A prospective study of hypofractionated proton beam therapy for patients with hepatocellular carcinoma. *Int. J. Radiat. Oncol. Biol. Phys.* 74, 831-836 (2009).
- Mizumoto M, Okumura T, Hashimoto T *et al.* Proton beam therapy for hepatocellular carcinoma: A comparison of three treatment protocols. *Int. J. Radiat. Oncol. Biol. Phys.* 81, 1039-1045 (2011).
- Qi WX, Fu S, Zhang Q *et al.* Charged particle therapy versus photon therapy for patients with hepatocellular carcinoma: A systematic review and meta-analysis. *Radiother. Oncol.* 114, 289-295 (2015).
- Hashimoto T, Tokuyue K, Fukumitsu N *et al.* Repeated proton beam therapy for hepatocellular carcinoma. *Int. J. Radiat. Oncol. Biol. Phys.* 65, 196-202 (2006).
- Oshiro Y, Mizumoto M, Okumura T *et al.* Analysis of repeated proton beam therapy for patients with hepatocellular carcinoma. *Radiother. Oncol.* 123, 240-245 (2017).
- Wognum S, Heethuis SE, Rosario T *et al.* Validation of deformable image registration algorithms on CT images of *ex vivo* porcine bladders with fiducial markers. *Med. Phys.* 41, 22-26 (2014).
- Yang D, Chaudhari SR, Goddu SM *et al.* Deformable registration of abdominal kilovoltage treatment planning CT and tomotherapy daily megavoltage CT for treatment adaptation. *Med. Phys.* 36, 329-338 (2009).
- Wang H, Dong L, Lii MF *et al.* Implementation and validation of a three-dimensional deformable registration algorithm for targeted prostate cancer radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 61, 725-735 (2005).
- Wen N, Glide HC, Nurushev T *et al.* Evaluation of the deformation and corresponding dosimetric implications in prostate cancer treatment. *Phys. Med. Biol.* 57, 5361-5379 (2012).
- Vasquez OEM, Hoogeman MS, Bondar L *et al.* A novel flexible framework with automatic feature correspondence optimization for nonrigid registration in radiotherapy. *Med. Phys.* 36, 2848-2859 (2009).
- Kirby N, Chuang C, Ueda U *et al.* The need for application-based adaptation of deformable image registration. *Med. Phys.* 40, 12-16 (2013).
- Moriya S, Tachibana H, Kitamura N *et al.* Dose warping performance in deformable image registration in lung. *Phys. Med.* 37, 16-23 (2017).
- Fukumitsu N, Nitta K, Terunuma T *et al.* Registration error of the liver CT using deformable image registration of MIM Maestro and Velocity AI. *BMC. Med. Imaging.* 17, 30 (2017).
- Nie K, Chuang C, Kirby N *et al.* Site-specific deformable imaging registration algorithm selection using patient-based simulated deformations. *Med. Phys.* 40, 14-18 (2013).
- Singhrao K, Kirby N, Pouliot J. A three-dimensional head-and-neck phantom for validation of multimodality deformable image registration for adaptive radiotherapy. *Med. Phys.* 41, 121-129 (2014).
- Kadoya N, Fujita Y, Katsuta Y *et al.* Evaluation of various deformable image registration algorithms for thoracic images. *J. Radiat. Res.* 55, 175-182 (2014).