Physica Medica xxx (2015) 1–5

Contents lists available at ScienceDirect

Physica Medica

journal homepage: http://www.physicamedica.com

Original paper

Validation of a deformable image registration produced by a commercial treatment planning system in head and neck

Rafael García-Mollá [a, *], Noelia de Marco-Blancas [a], Jorge Bonaque [a], Laura Vidueira [a], Juan López-Tarjuelo [a], José Perez-Calatayud [b]

[a] Medical Physics and Radioprotection Department, Consorcio Hospitalario Provincial de Castellón, C/ Dr. Clara, nº 19, Castellón de la Plana, 12002, Spain
[b] Radiotherapy Department, La Fe Polytechnic and University Hospital, Bulevar Sur s/n, Valencia, 46026, Spain

ARTICLE INFO

Article history:
Received 29 July 2014
Received in revised form
14 January 2015
Accepted 16 January 2015
Available online xxx

Keywords:
Deformable image registration
Adaptive radiotherapy
IMRT

ABSTRACT

In recent years one of the areas of interest in radiotherapy has been adaptive radiation therapy (ART), with the most efficient way of performing ART being the use of deformable image registration (DIR). In this paper we use the distances between points of interest (POIs) in the computed tomography (CT) and the cone beam computed tomography (CBCT) acquisition images and the inverse consistence (IC) property to validate the RayStation treatment planning system (TPS) DIR algorithm. This study was divided into two parts: Firstly the distance-accuracy of the TPS DIR algorithm was ascertained by placing POIs on anatomical features in the CT and CBCT images from five head and neck cancer patients. Secondly, a method was developed for studying the implication of these distances on the dose by using the IC. This method compared the dose received by the structures in the CT, and the structures that were quadruply-deformed. The accuracy of the TPS was 1.7 ± 0.8 mm, and the distance obtained with the quadruply-deformed IC method was 1.7 ± 0.9 mm, i.e. the difference between the IC method multiplied by two, and that of the TPS validation method, was negligible. Moreover, the IC method shows very little variation in the dose-volume histograms when comparing the original and quadruply-deformed structures. This indicates that this algorithm is useful for planning adaptive radiation treatments using CBCT in head and neck cancer patients, although these variations must be taken into account when making a clinical decision to adapt a treatment plan.

Introduction

One of the areas of greatest interest in radiotherapy is adaptive radiation therapy (ART), with deformable image registration (DIR) being the most efficient way of performing ART [1]. In order to clinically adapt a treatment plan the DIR must be validated, and so researchers and users must determine whether any dose changes observed are due to poor DIR algorithm accuracy or due to anatomical changes in the patient.

There are different methods to test DIR algorithms, including the use of: (a) landmarks or/and contours in two separate acquisition images from a patient's computed tomography (CT) scan or cone beam computed tomography (CBCT) [2,3], (b) physical and/or deformable dosimetric phantoms [4–7], or (c) dedicated software applications. [8–10], and each method has specific limitations. For instance, specific quality assurance (QA) software does not take image acquisition into account and the problem when using contours or landmarks is that point selection or structure creation is uncertain. Additionally, it must be added that only a few voxels are studied [2]. One of the general limitations of these methods is that all of them, except for dose deformable phantoms, give results in terms of distances whereas absorbed dose quantity is of most interest in radiotherapy [11]. Moreover, many hospitals do not have access to physical phantoms.

The distance between landmarks is widely used to calculate the accuracy of DIR algorithms [2,3], and the use of the inverse consistency (IC) property as a validation method has also been previously studied [12–14]. The IC property is very desirable in DIR algorithms, and validation studies using IC properties have shown that their value greatly depends on the specific DIR algorithm used [13,14]. For instance, suppose that image A is first deformed into image B and then B is deformed into A by an algorithm with an IC property. Consequently, this algorithm would have the same deformation vector fields (DVF) but these would be inverse to each

* Corresponding author.
E-mail address: rafael.garcia@hospitalprovincial.es (R. García-Mollá).

http://dx.doi.org/10.1016/j.ejmp.2015.01.007
1120-1797/

other. In this case, if the structures/regions of interest (ROIs) are mapped from image A to image B using the first DVF (DVF_1) and later mapped back from B to A with the second DVF (DVF_2), both the original and the doubly-deformed structures should theoretically be equal.

In this study, we use the distances between points of interest (POIs) in the CT and CBCT acquisition images and the IC property to validate a the RayStation's treatment planning system (TPS) DIR algorithm (v.4.0.1.4, RaySearch Laboratories AB, Stockholm, Sweden) [12]. This study was divided into two parts: firstly, the distance-accuracy of the TPS hybrid DIR algorithm was ascertained by placing POIs on anatomical features in CT and CBCT acquisition images obtained for head and neck cancer patients, and the distances from these POIs mapped from the CBCT to CT were measured. Secondly, a new method was developed for studying the implication these distances have on the absorbed dose in distinct anatomical regions, by using the IC property of the DIR algorithm to compare the absorbed dose received by the structures contoured in the CT by the physician, and the structures which were quadruply-deformed by the hybrid DIR algorithm.

Materials and methods

DIR algorithms

The RayStation TPS has two DIR algorithms: the first one is the hybrid DIR which is based on a mathematical formula in which the registration is a non-linear optimization problem. The objective function is composed of four terms: 1) to maintain image similarity, 2) to keep the image grid smooth and invertible, 3) to keep the deformation anatomically reasonable when structures are present, and 4) a penalty term when structures are used. The second algorithm is the structure-based DIR that is based entirely on controlling the ROI and POI regions which are created within the planning CT (pCT) and CBCT. This study is focussed on validating the hybrid DIR algorithm because our aim was to validate the use of adaptive radiation using CBCT in an efficient way. It should be noted that although the invertibility condition of the vector field is assured with the hybrid RayStation algorithm (term 2 of the objective function), this does not mean that the IC property is met.

Patient anatomic data

For this study we selected five head and neck cancer patients. These patients were treated with step and shoot intensity-modulated radiation therapy (IMRT) and were positioned using an Elekta Synergy™ XVI image-guided radiotherapy-capable linear accelerator kV-CBCT imaging system (Crawley, UK; release 4.2.1). For each patient, pCT and three CBCT images were used to validate the DIR. The CBCT images were taken in the first and last weeks and during a week in the middle of the treatment to take into account the influence of possible anatomical changes in the patients on the accuracy of the DIR. The pCT images were acquired with a Siemens SOMATOM Sensation 16 CT scanner (Siemens AG, Erlangen, Germany), with a slice thickness of 3 mm and a pixel size of 1 mm, and were acquired using the previously mentioned XVI CBCT imager with a slice thickness and pixel size of 1 mm.

The deformable image registration process

The process of comparing planned dose vs. delivered accumulated absorbed dose with the TPS RayStation consists of three stages: 1) A rigid registration is performed with pCT and CBCT images; the DIR is then performed and thus the ROIs/POIs in the two registered images can be mapped in either direction (the CBCT to pCT or vice versa). 2) After the doses are calculated for different CBCT images, they are projected to the pCT by the DVF. 3) The plan is then adapted by comparing this information with the planned pCT dosimetry and the DIR dosimetry accumulated from the CBCTs.

Distance validation

To estimate the accuracy of the DIR distances, a radiation oncologist selected ten anatomically recognizable features in the pCT acquisition image (reference POIs) from the five head and neck cancer patients with the POI tool (Fig. 1), and selected them again in the same areas in the three CBCTs which were acquired. Half of the POIs were selected in soft areas and the other half were in rigid areas. Table 1 shows the areas where the POIs were placed. To reduce potential variation in operator POI position selection and to assess the observer variability, three different observers selected the same points on each of the CBCT and pCT images, using hard copies of pCT images as a reference. The accuracy of the hybrid DIR algorithm was quantified using the distance between the POIs mapped from CBCT to pCT and the POIs of the pCT. DIRs were performed using a grid size of 2.5 mm as recommended by the manufacturer and the contour from a skin patient was used as a control ROI. This external contour was used as a controlling ROI to force the deformable registration to focus more on matching the outer contour of the patient. The observer variability was quantified using the distance between the POIs of the radiation oncologist and the average POIs of the observers.

The four points per location (created by the three observers and radiation oncologists in pCT and in each of the CBCTs) were averaged to reduce inter-observer point placement variation and the average POI was mapped from the CBCT to the pCT using the DVF_1. The difference in the distances between the each POI projected from the CBCT and its corresponding reference POI determined the DIR distance accuracy. The TPS gave the coordinates of the POIs with a resolution of 0.1 mm. Six hundred (600) POIs were placed on the CBCT images to validate the DIR algorithm, and two hundred (200) were placed on the pCT images of the five patients to evaluate the observer variability.

To discard any suspected outliers in the observer data (the POI selected by the observer), suspicious results are discarded when the standard deviation of this value is at least four times the average standard deviation of the other results. Four times the average standard deviation is an arbitrary value which is commonly used for detecting outliers and is very unlikely to exclude any valid information.

Inverse consistency method validation

The purpose of the IC method is to determine the accuracy in dose values of the algorithm. As the hybrid algorithm does not have a perfect IC, the original and deformed structures (projected from the pCT to the CBCT followed by projection from the CBCT back to the pCT) do not coincide on the pCT image and therefore a difference is observed in the HDV. In this way, we can take advantage of the lack of IC to determine the dose accuracy of the hybrid algorithm in terms of doses rather than distances. In order to validate this method, we must ensure that the distances obtained using the IC method coincides with those obtained in the previous section since we assume that this distance represents the accuracy of the algorithm.

The process used to determine the accuracy of the IC in distances is as follows: the pCT POIs are projected from the pCT to the CBCT by DVF_2 and then back from CBCT to the pCT by DVF_1. To obtain distances comparable to the IC method and the accuracy of the algorithm, the process was performed twice, meaning that the

Please cite this article in press as: García-Mollá R, et al., Validation of a deformable image registration produced by a commercial treatment planning system in head and neck, Physica Medica (2015), http://dx.doi.org/10.1016/j.ejmp.2015.01.007

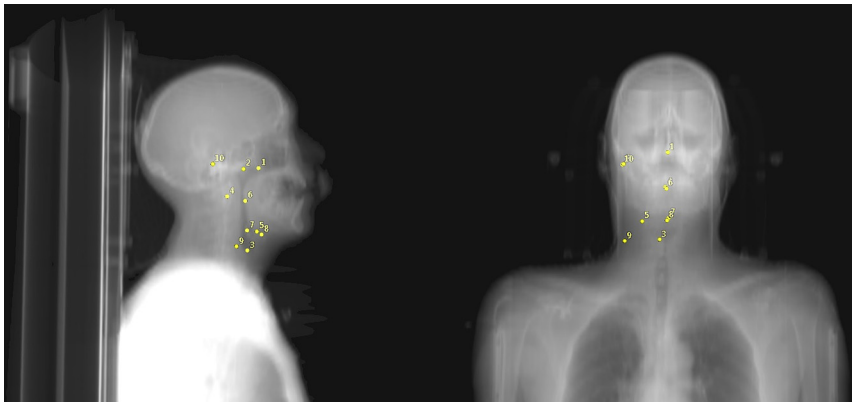

Figure 1. Points-of-interest (POIs) are represented by yellow spheres in the planning computed tomography (pCT) images from a head and neck region. Table 1 shows the anatomical positions of POIs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

POIs are deformed four times (quadruply-deformed points). The POIs are projected from the pCT to the CBCT, then back from the CBCT to the pCT, and the process is repeated again (pCT to CBCT and CBCT to pCT). In DVF_1, the CBCT is selected as reference image and the pCT is deformed; for DVF_2, the pCT is selected as reference image and the CBCT is deformed. Exactly the same process is carried out with the structures contoured by the physician in the pCT and so, if the IC method distances and validation distances are comparable, IC method dose-validation is correct. Figure 2 shows a diagram explaining these two processes.

Dose validation

Although it is important to determine the DIR distance accurately, the most important quantity in radiotherapy is absorbed dose. Uncertainty in DIR distance has different effects on the absorbed dose, depending on the treatment zone [9]. The procedure used to evaluate the dose accuracy of the hybrid DIRs algorithm is based on the IC metric: structures contoured in the pCT are projected onto the CBCT by DVF_2 and then these structures are projected from the CBCT to the original pCT by DVF_1. As for the POIs (refer to 2.5 section), this process is performed twice, so, in total, the structures are deformed four times. This quadruple-deformation method eliminates the influence of CBCT dose-calculation in the calculation of DIR accuracy [15]. The size of the field of view (FOV) of CBCT must be taken into account when comparing the original and the quadruply-deformed structures as these structures can be out of FOV. These structures were not included in the study.

Table 1
Anatomical positions of the ten POIs placed in each patient's pCT and CBCT.

POI	Anatomical areas
1	Nasal septum
2	Right temporomandibular Joint
3	Right vocal cord
4	Odontoid apophysis
5	Right submandibular gland
6	Uvula
7	Epiglottis
8	Hyoid bone
9	Right sternocleidomastoid muscle
10	Right mastoid air cells

Results and discussion

Distance validation

Table 2 shows the average distance measured of the POIs projected from the CBCT to the pCT, and of the POIs into the acquisition pCT, and the *SD* (coverage factor $k = 1$) for the five head and neck cancer patients included in the study. Figure 3 shows the accuracy of DIR for each area. The POIs in soft and rigid areas had an average distance of 2.0 ± 0.1 and 1.4 ± 0.5 mm respectively. The average distance was 1.7 ± 0.8 mm, which equals or is less than the accuracy of other algorithms [2,3]. Only ten POIs had an average distance of more than 3 mm, and using the 4*SD* criterion to discard outliers, 17 (2.8%) of 600 placed POIs were discarded. The average distance of observer variability was 0.06 ± 0.04 mm.

Inverse consistency method validation

Table 3 shows the distances from the averaged original pCT POIs, and the doubly-deformed and quadruply-deformed POIs. The distances obtained with the IC method for doubly-deformed POIs

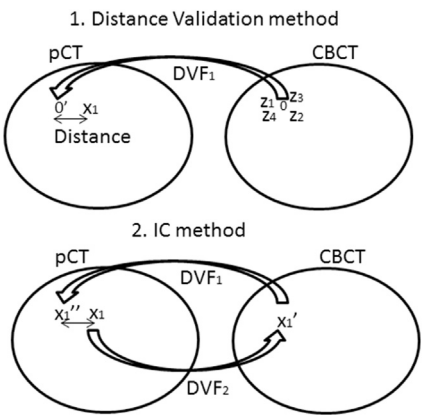

Figure 2. Top of the diagram (1) shows the validation distances method: The radiation oncologist selected the same POI in the pCT (x1) and in the cone beam computed tomography (CBCT; z1) and then three observers selected the same POI in the CBCT (z2, z3, z4). The four CBCT points are averaged (0), and this point is projected by the second deformation vector field (DVF_2) to the pCT (0′). The distances between 0′ and x1 represent the accuracy of the deformable image registration (DIR) algorithm. The bottom of the diagram (2) shows the inverse consistency (IC) method: The point selected by the radiation therapist is projected from pCT to CBCT by DVF_1, and after from CBCT to the pCT by DVF_2. The distances between x1 and x1″ gives the IC metric result from the DIR algorithm.

Please cite this article in press as: García-Mollá R, et al., Validation of a deformable image registration produced by a commercial treatment planning system in head and neck, Physica Medica (2015), http://dx.doi.org/10.1016/j.ejmp.2015.01.007

Table 2
Average distance and standard deviation (coverage factor k = 1) of the POI DIRs projected from the CBCT to the TPS and pCT POIs. The POIs were projected using the RayStation TPS hybrid algorithm for the five different head and neck cancer patients included in the study.

Patient	POI distance (mm)
1	1.6 ± 0.8
2	1.8 ± 0.9
3	1.5 ± 0.7
4	1.7 ± 0.7
5	1.8 ± 0.6
Average	1.7 ± 0.7

Table 3
Distances from the averaged original pCTs to the doubly-deformed and quadruply-deformed POIs in the five different head and neck cancer patients included in the study.

Patient	Distance	
	Doubly-deformed POIs (mm)	Quadruply-deformed POIs (mm)
1	0.7 ± 0.3	1.4 ± 0.5
2	0.7 ± 0.4	1.5 ± 0.7
3	0.8 ± 0.4	1.4 ± 0.6
4	0.9 ± 0.5	1.8 ± 1.0
5	1.3 ± 0.7	2.5 ± 1.4
Average	0.9 ± 0.4	1.7 ± 0.9

(0.9 mm) are close to half of the validation distance (1.7 mm). This indicates that the uncertainty of the IC of this algorithm is lower than the uncertainty of the DIR algorithm. As the difference in distances was nearly half of the validation distances we decided to repeat the IC method (quadruply-deformed) to ensure that the distances were similar. The distance obtained with the quadruply-deformed IC method was 1.7 ± 0.9 mm, comparable with the DIR validation distance of 1.7 ± 0.8 mm.

Dose validation

Table 4 shows differences in the DVHs, between the original pCT structures and the quadruply-deformed structures. There was more DVH variation in the temporomandibular joint (TMJ) than in other structures due to its small size, meaning that any small DIR error greatly impacted the TMJ dose-value. Moreover, the parotid, the brainstem, and the brain are poor-contrast areas, thus reducing the accuracy of the DIR produced by the TPS in these areas. These have a high dose-gradient which means that small differences in the distances result in large dose-differences delivered (over 1 Gy). In the case of spinal cord, small differences in volume have a great impact in the average dose and the D_{50}. When the average dose value differences are high (greater than 1Gy), the *SD* is also important, indicating that there is great variation among the differences. However, even taking these issues into account, there was very little DVH variation compared with the absolute dose values, indicating that the DIR algorithm works reasonably well. This indicates, for the first time in head and neck region cancer patients, that the hybrid DIR algorithm is useful for planning adaptive radiation treatments using CBCT, although radiation

Table 4
Differences in the DVHs, i.e. the dose-value, between the original pCT structures and the quadruply-deformed (IC method) structures. TMJ: temporomandibular joint; CTV: clinical target volume; PTV: planning target volume.

	Volume (cm^3)	Differences				
		D_{98}	D_{95}	Average	D_{50}	D_2
		(cGy)				
TMJ	0.18 ± 0.05	50 ± 30	50 ± 40	99 ± 70	130 ± 10	100 ± 70
Brain stem	11 ± 8	40 ± 20	50 ± 20	100 ± 100	50 ± 60	90 ± 50
Brain	40 ± 50	0.2 ± 0.2	0.1 ± 0.2	20 ± 20	20 ± 30	200 ± 100
Larynx	1.1 ± 0.5	80 ± 90	80 ± 80	50 ± 10	60 ± 40	100 ± 100
Jaw	12 ± 1	70 ± 70	60 ± 50	40 ± 30	50 ± 30	34 ± 03
Cord	2 ± 1	70 ± 90	60 ± 60	110 ± 80	200 ± 300	52 ± 70
Optic nerve	0.24 ± 0.01	30 ± 10	30 ± 10	36 ± 2	30 ± 3	60 ± 30
Eyes	7 ± 10	20 ± 20	30 ± 20	20 ± 20	30 ± 20	45 ± 30
Parotid	1.4 ± 0.6	60 ± 50	50 ± 30	80 ± 40	70 ± 40	100 ± 200
PTV	10 ± 9	60 ± 60	40 ± 30	40 ± 40	20 ± 40	30 ± 60
CTV	6 ± 6	30 ± 40	30 ± 40	40 ± 80	20 ± 30	40 ± 70

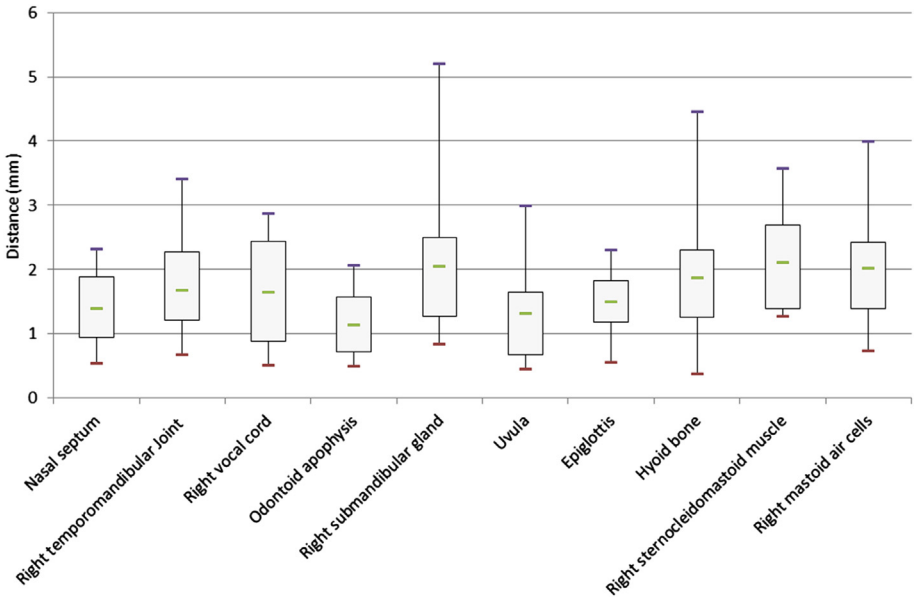

Figure 3. Box-and-whisker plot showing the accuracy of each area. The points represent the maximum, the upper quartile, the mean, the lower quartile and the minimum distance.

Please cite this article in press as: García-Mollá R, et al., Validation of a deformable image registration produced by a commercial treatment planning system in head and neck, Physica Medica (2015), http://dx.doi.org/10.1016/j.ejmp.2015.01.007

oncologists should remain cautious when adapting clinical plans, carefully taking the differences in DVH into consideration.

Conclusions

The accuracy of TPS hybrid DIR algorithm was ascertained by placing POIs on anatomical landmarks in the CT and CBCT. The distance from these POIs was 1.7 mm.

A new method was developed to study the implication of these distances on the absorbed dose, without a physical phantom or specific software, by using the IC property twice (quadruply-deformed structures). This method has been validated by comparing the distances obtained using the IC method with those from the DIR validation method; we found the difference to be negligible. Although the IC method allows to validate in doses the hybrid deformable registration algorithm without a physical phantom, these are considered essential for validating DIR algorithms. In future work we hope to show our progress on this area.

The TPS hybrid DIR algorithm shows small amounts of variation in the DVH. This indicates that this algorithm is useful for adaptive radiation treatment planning using CBCT for head and neck cancer patients, although these variations must be taken into account before making any clinical decisions to adapt a plan.

Acknowledgments

The authors wish to thank Stina Svensson (RaySearch) and Facundo Ballester (University of Valencia) for critical reading of the manuscript and Maria Hannah Ledran and the Fundación Hospital Provincial de Castellón, for the English manuscript revision. This study was supported in part by Spanish Government under Project No. FIS2013-42156. The authors declare no conflicts of interest.

References

[1] Wang H, Dong L, Lii MF, Lee AL, Crevoisier RD, Mohan R, et al. Implementation and validation of a three-dimensional deformable registration algorithm for targeted prostate cancer radiotherapy. Int J Radiat Oncol Biol Phys 2005;61:725–35.
[2] Mencarelli A, van Beek S, van Kranen S, Rasch C, van Herk M, Sonke JJ. Validation of deformable registration in head and neck cancer using analysis of variance. Med Phys 2012;39:6879–84.
[3] Huger S, Graff P, Harter V, Marchesi V, Royer P, Diaz JC, et al. Evaluation of the block matching deformable registration algorithm in the field of head-and-neck adaptive radiotherapy. Phys Medica 2014;30:301–8.
[4] Kirby N, Chuang C, Ueda U, Pouliot J. The need for application-based adaptation of deformable image registration. Med Phys 2013;40. 011702–1-10.
[5] Kashani R, Hub M, Kessler ML, Balter JM. Technical note: a physical phantom for assessment of accuracy of deformable alignment algorithms. Med Phys 2007;34:2785–8.
[6] Yeo UJ, Taylor ML, Supple JR, Smith RL, Dunn L, Kron T, et al. Is it sensible to "deform" dose? 3D experimental validation of dose-warping. Med Phys 2012;39:5065–72.
[7] Yeo UJ, Taylor ML, Dunn L, Kron T, Smith RL, Franich RD. A novel methodology for 3D deformable dosimetry. Med Phys 2012;39:2203–12.
[8] Varadhan R, Karangelis G, Krishnan K, Hui S. A framework for deformable image registration validation in radiotherapy clinical applications. J Appl Clin Med Phys 2013;14:192–211.
[9] Zhong H, Kim J, Chetty J. Analysis of deformable image registration accuracy using computational modeling. Med Phys 2010;37:970–9.
[10] Nie K, Chuang C, Kirby N, Braunstein S, Pouliot J. Site-specific deformable imaging registration algorithm selection using patient-based simulated deformation. Med Phys 2013;40. 041911–1-10.
[11] Saleh-Sayah NK, Weiss E, Salguero FJ, Siebers JF. A distance to dose difference tool for estimating the required/spatial accuracy of a displacement vector field. Med Phys 2011;38:2318–23.
[12] Yang D, Li H, Low DA, Deasy JO, Naqa IE. A fast inverse consistent deformable image registration method based on symmetric optical flow computation. Phys Med Biol 2008;53:6143–65.
[13] Bender ET, Tomé WA. The utilization of consistency metrics for error analysis in deformable image registration. Phys Med Biol 2009;54:5561–77.
[14] Yan C, Zhong H, Murphy M, Weiss E, Siebers JV. A pseudoinverse deformation vector field generator and its applications. Med Phys 2010;37:1117–28.
[15] Seet KYT, Barghi A, Yartsev S, Van Dyk J. The effects of field-of-view and patient size on CT numbers from cone-beam computed tomography Phys. Med Biol 2009;54:6251–62.

Please cite this article in press as: García-Mollá R, et al., Validation of a deformable image registration produced by a commercial treatment planning system in head and neck, Physica Medica (2015), http://dx.doi.org/10.1016/j.ejmp.2015.01.007