NONRIGID REGISTRATION METHOD TO ASSESS REPRODUCIBILITY OF BREATH-HOLDING WITH ABC IN LUNG CANCER

DAVID SARRUT, PH.D.,*† VLAD BOLDEA,† MYRIAM AYADI,* JEAN-NOËL BADEL,* CHANTAL GINESTET,* SÉBASTIEN CLIPPE, M.D.,* AND CHRISTIAN CARRIE, M.D.*

*Department of Radiotherapy, Centre Léon Bérard, Lyon, France; †LIRIS Laboratory, Université Lumière Lyon, Lyon, France

Purpose: To study the interfraction reproducibility of breath-holding using active breath control (ABC), and to develop computerized tools to evaluate three-dimensional (3D) intrathoracic motion in each patient.

Methods and Materials: Since June 2002, 11 patients with non–small-cell lung cancer enrolled in a Phase II trial have undergone four CT scans: one during free-breathing (reference) and three using ABC. Patients left the room between breath-hold scans. The patient’s breath was held at the same predefined phase of the breathing cycle (about 70% of the vital capacity) using the ABC device, then patients received 3D-conformal radiotherapy. Automated computerized tools for breath-hold CT scans were developed to analyze lung and tumor interfraction residual motions with 3D nonrigid registration.

Results: All patients but one were safely treated with ABC for 7 weeks. For 6 patients, the lung volume differences were <5%. The mean 3D displacement inside the lungs was between 2.3 mm (SD 1.4) and 4 mm (SD 3.3), and the gross tumor volume residual motion was 0.9 mm (SD 0.4) to 5.9 mm (SD 0.7). The residual motion was slightly greater in the inferior part of the lung than the superior. For 2 patients, we detected volume changes >300 cm³ and displacements >10 mm, probably owing to atelectasia and emphysema. One patient was excluded, and two others had incomplete data sets.

Conclusion: Breath-holding with ABC was effective in 6 patients, and discrepancies were clinically accountable in 2. The proposed 3D nonrigid registration method allows for personalized evaluation of breath-holding reproducibility with ABC. It will be used to adapt the patient-specific internal margins. © 2005 Elsevier Inc.

Breath-holding, Nonrigid registration, Lung cancer.

INTRODUCTION

Organ motion due to breathing is a source of inaccuracy in treatment delivery because it leads to tumor displacement and suboptimal dose delivery. The three-dimensional (3D) motion of lung tumors has been studied by several authors (1–9), revealing its complexity. In particular, it has been studied with gold markers and fluoroscopy (1, 2). The tumor motion amplitude varies with localization and can reach 12 mm (lower lobes, near the diaphragm). A main challenge for non–small-cell lung cancer (NSCLC) radiotherapy (RT) is to spare the surrounding normal tissues while providing the prescribed doses to the tumor, because although dose escalation seems to yield superior outcomes (3, 10, 11), lung complications have been shown to correlate with the dose delivered to the normal lung (12–15).

Limiting organ motion can be achieved using multiple approaches (16), including patient immobilization, breath control on instruction, abdominal compression, deep inspiration breath hold (DIBH), active breath control (ABC), respiratory gating, target tracking (listed roughly from simple to more complex). Patient immobilization is commonly used and has proved effective (17, 18) in reducing motion. Abdominal compression or breath coaching also reduces motion (18). However, these techniques do not totally eliminate motion. Other techniques try to immobilize the organs and tumors with breath-holding (19). Patients generally hold their breath at deep inspiration (DIBH), which increases the lung volume and consequently decreases the mass of irradiated lung. It can be implemented passively (patient holds breath on instruction) or actively (with the so-called ABC device). Respiratory gating is more comfortable for patients. It consists of synchronizing radiation delivery with patient breathing. The treatment beam is enabled at predetermined intervals during the breathing cycle while the patient breathes freely (4, 9, 19–22). This technique depends on a correlation between external and internal movement, which has not yet been well established (4) and requires extensive physics support. Target tracking is the most ambitious strategy (23). It consists of continuously adapting the radiation beam delivery to changes in the tumor position. It requires tremendous effort and remains a challenge.

Reprint requests to: David Sarrut, Ph.D., Department of Radiotherapy, Centre Léon Bérard, 28 rue Laennec, Lyon 69008 France; Tel: (+33) 4-7878-5151; Fax: (+33) 4-7878-2626; E-mail: dsarrut@univ-lyon2.fr

The next step after efforts to limit organ motion was to quantify the residual motion and adapt the treatment (16). We were interested in immobilizing the organs and tumors by active breath holding, so we only reviewed the literature published in this field. Organ motion is generally decomposed into intrafraction and interfraction motion (24). For breath-hold methods, we must also differentiate between intra-breath-hold motion, which is the residual movement that could possibly occur during a breath hold, and intrafraction motion that occurs between two breath holds without patient repositioning. Intra-breath-hold motion was studied by Dawson et al. (25, 26) with videotaped fluoroscopy. No motion of the diaphragm or fiducial markers was observed. The mean intrafraction motion was estimated at between 1 and 2.5 mm with active breath-holding (17, 26, 27). Dawson and Balter (25) observed that 90% of the differences in marker positions were <4.8 mm. Interfraction reproducibility varies with the phases of the respiratory cycle: it tended to be better at the end of normal expiration/inspiration phases than at deep inspiration for some authors (26, 28), although contrary findings have also been reported (29). Details on interfraction reproducibility are given in Table 1.

Table 1. Review of interfraction reproducibility studies

<table>
<thead>
<tr>
<th>Dim</th>
<th>Reference</th>
<th>Intrafraction difference mm (SD)</th>
<th>Features</th>
<th>A/P</th>
<th>Level</th>
<th>Site</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D</td>
<td>Remouchamps et al. (31), 2003</td>
<td>CC = 3.2; ML = 2.4; rot = 1°</td>
<td>Field border (medium beam)</td>
<td>ABC</td>
<td>mDIBH</td>
<td>Breast</td>
<td>EPID (field border), skin tattoo for setup</td>
</tr>
<tr>
<td>2D</td>
<td>Balter et al., (34), 2002</td>
<td>CC = 3.1; ML = 2.3; rot = 1°</td>
<td>Field border (lateral beam)</td>
<td>ABC</td>
<td>NE</td>
<td>Liver</td>
<td>CT/film comparisons Radiographs Radiographs, microcoils</td>
</tr>
<tr>
<td>2D</td>
<td>Dawson et al., (26), 2001</td>
<td>AP = 2.3; ML = 2.1; CC = 2.5</td>
<td>Skeleton Diaphragm</td>
<td>ABC</td>
<td>NE</td>
<td>Liver</td>
<td>Fluoroscopy, film, video</td>
</tr>
<tr>
<td>2D</td>
<td>Kim et al., (29), 2001</td>
<td>CC &lt; 5</td>
<td>Diaphragm</td>
<td>Self</td>
<td>NE/NI DE/DI</td>
<td>NSCLC</td>
<td>Portal films</td>
</tr>
<tr>
<td>2D</td>
<td>Mah et al., (53), 2000</td>
<td>−12 to 11/1.3 (5.3) range 9.1–0.2 (1.4)</td>
<td>Diaphragm GTV center</td>
<td>Self</td>
<td>DI</td>
<td>NSCLC</td>
<td>Fluoroscopy</td>
</tr>
<tr>
<td>2D</td>
<td>Hanley et al., (54), 1999</td>
<td>2.5 (1.6)</td>
<td>Diaphragm</td>
<td>Self</td>
<td>DI</td>
<td>NSCLC</td>
<td>5 CT</td>
</tr>
<tr>
<td>3D</td>
<td>Cheung et al., (11), 2003</td>
<td>CC = 1.1 (3.5); AP = 1.2 (2.3); ML = 0.3 (1.8)</td>
<td>GTV center</td>
<td>ABC</td>
<td>DI</td>
<td>NSCLC</td>
<td>3 CT</td>
</tr>
<tr>
<td>3D</td>
<td>Onishi et al., (55), 2003</td>
<td>CC = 2.2 (1.1); AP = 1.4 (0.6); ML = 1.3 (0.5)</td>
<td>Tumor</td>
<td>Self</td>
<td>DI</td>
<td>NSCLC</td>
<td>5 CT</td>
</tr>
<tr>
<td>3D</td>
<td>Remouchamps et al., (32), 2003</td>
<td>1.4 (1.7); 1.4 (1.0); 1.9 (2.2)</td>
<td>Lung surface Trachea Diaphragm</td>
<td>ABC</td>
<td>mDIBH</td>
<td>Breast</td>
<td>2 CT, DTA</td>
</tr>
<tr>
<td>3D</td>
<td>Onishi et al., (55), 2003</td>
<td>CC = 2.1; AP = 1.4; ML = 1.3</td>
<td>Tumor</td>
<td>Self</td>
<td>DI</td>
<td>Lung</td>
<td>3 CT (active mode)</td>
</tr>
<tr>
<td>3D</td>
<td>Wilson et al., (27), 2003</td>
<td>CC = 3.1; AP = 2.4; ML = 2.2</td>
<td>Lung volume</td>
<td>ABC</td>
<td>mDIBH</td>
<td>NSCLC</td>
<td>3 CT</td>
</tr>
<tr>
<td>3D</td>
<td>Stromberg et al., (36), 2000</td>
<td>0.2 to 8.7%, max = 13.2%</td>
<td>Lung volume</td>
<td>ABC</td>
<td>DI</td>
<td>Hodgkin’s</td>
<td>5 CT</td>
</tr>
<tr>
<td>3D</td>
<td>Wong et al., (30), 1999</td>
<td>4.0%</td>
<td>Lung volume</td>
<td>ABC</td>
<td>NE/DI</td>
<td>Lung, liver, Hodgkin’s</td>
<td>2 CT</td>
</tr>
<tr>
<td>3D</td>
<td>Hanley et al., (54), 1999</td>
<td>0.2 to 8.7%, max = 13.2%</td>
<td>Lung volume</td>
<td>ABC</td>
<td>mDIBH</td>
<td>NSCLC</td>
<td>3 CT</td>
</tr>
<tr>
<td>3D</td>
<td>Wilson et al., (27), 2003</td>
<td>4.0%</td>
<td>Lung volume</td>
<td>ABC</td>
<td>DI</td>
<td>Hodgkin’s</td>
<td>5 CT</td>
</tr>
<tr>
<td>3D</td>
<td>Stromberg et al., (36), 2000</td>
<td>6.0%</td>
<td>Lung volume</td>
<td>ABC</td>
<td>NE/DI</td>
<td>Lung, liver, Hodgkin’s</td>
<td>2 CT</td>
</tr>
<tr>
<td>3D</td>
<td>Hanley et al., (54), 1999</td>
<td>Lat width = 1.1%; AP height = 1.5%; lung area = 3%</td>
<td>Overlapping slices</td>
<td>Self</td>
<td>DI</td>
<td>NSCLC</td>
<td>2–4 CT</td>
</tr>
</tbody>
</table>

Abbreviations: A/P = active/passive breath hold; Dim = dimension; 2D = two-dimensional; rot = rotation; Lat = lateral; CC = Craniocaudal; ML = Medial–lateral; ABC = active breath control; DIBH = deep inspiration breath hold; mDIBH = medium DIBH; NE = near end of normal expiration; NI = near end of normal inspiration; DE = deep expiration; EPID = electronic portal imaging device; DTA = distance to agreement.
The goal of this study was to develop a nonrigid (deformable) 3D registration method to evaluate systematically the interfraction reproducibility of breath-hold with the ABC device. This approach estimates the 3D displacement of each point between different CT scans. This helps quantify tumor residual motion using the ABC device and, in future studies, will make it possible to personalize internal margins in treatment planning.

METHODS AND MATERIALS

Data acquisition

Patients. Since June 2002, 11 patients with NSCLC have been enrolled in a Phase II trial. Radical RT was indicated because of potentially resectable but inoperable T1–T4, N0–N1, M0 NSCLC. All patients had severe respiratory insufficiency. All patients provided informed consent in accordance with French laws and with the procedures of the local Consultative Committee for the Protection of Participants in Biomedical Research.

ABC device. The ABC device proposed by Wong et al. (30) allows for temporary immobilization of respiratory motion by implementing a breath hold at a predefined lung volume (relative to the end of normal expiration, corresponding to the functional residual capacity) and air flow direction. When the patient reaches this predefined lung volume level during either inspiration or expiration, the airflow is temporarily blocked, thereby immobilizing breathing motion. The radiation source is turned on during this period. To our knowledge, several studies have reported the use of ABC for liver (26), breast (17, 31–35), and lung (11, 27, 30, 36) cancers.

Patient breath-holding with ABC. We used two ABC devices from Elekta (Active Breathing Coordinator), one in the treatment room, the other in the CT scan room. The two devices were calibrated with a 3-L syringe. The patient’s breath was held at approximately the same given phase of the breathing cycle corresponding to about 70%/75% (or medium DIBH [17]) of the vital capacity; other studies have used 80% (30) or 75% (17, 27). The breath-hold time was about 20 s, as in other studies (11, 26, 27, 30), except that the time was extended to 30–45 s for some patients with Hodgkin’s cancer (30, 36) or liver cancer (34). All patients underwent an initial training session to determine the maximal time gate they could reach comfortably. This is particularly important because they had to repeat this step as many times as the total number of fractions required. The training session and preparation of patient immobilization took approximately 90 mins. Verbal instructions were given all along the sessions. The breath-hold duration was 15–20 s, although additional time could sometimes be allocated, depending on the patient’s ability. The initial sessions were generally longer than those following, because of the patient’s habituation. Patient immobilization was achieved using an α-craddle device (17).

CT scans. Each patient underwent one CT scan during free breathing (reference) and three in breath-hold mode using the ABC device, at the same breath-hold level for each scan. All CT scans were acquired on a Picker PQ 2000 scanner (Picker International Inc., Tampa, FL). Free-breathing scans were acquired in axial mode and breath-hold scans in spiral mode (spiral pitch of 1.5). Free-breathing scans were acquired in time with an average respiratory cycle, about 4 s (3) to get a blurred tumor image that supposedly covered the tumor motion range (37). Breath-hold scans were acquired over several breath-hold sessions (seven or eight depending on patient ability), as described by Cheung et al. (11). Slices (5 mm thickness) were then stacked into a 3D volume. Consistency between consecutive slices was checked by visual inspection. The total acquisition session took about 45 min for the one breath-hold CT scan. Patients left the room after each breath-hold acquisition, allowing interfraction comparisons. The final 3D data set size range was 512 × 512 with 60–70 slices. Eight data sets of 11 were fully available (1 patient was excluded and 2 had incomplete data sets owing to a storage failure), with 24 ABC-aided CT scans and 8 free breathing scans.

Treatment. Patients then underwent 3D-conformal RT. The treatment consisted of two steps: one step using six coplanar static fields with a multileaf colimator (MLC) for a dose of up to 40–50 Gy and the second using six boost fields and a MLC to a final dose of 70 Gy (using the Elekta SL20 accelerator). Each field was delivered with patients holding their breath using the ABC device at the predetermined threshold fitted to the lung dose–volume histogram. Patients received 35 fractions (one fraction daily) with six breath holds for each fraction, for a total of >200 breath holds using ABC. The additional time for each session was between 5 and 10 min. This is consistent with other reports, one of which (26) reported an approximate extra time of 10 min, and one (17) that reported a treatment duration of <20 min.

Lung volume comparison

Our first goal was to develop a reproducible tool for automatic quantification of lung volumes (left/right/whole lung). The first step consisted of segmenting the whole lung with simple thresholding (27, 30, 36). Voxels with a density below a given threshold were selected. These voxels corresponded to the air (outside and inside the patient), lung, and gas (e.g., in the patient’s bowel). Then, the resulting binary image was labeled with an automated 3D connected components labeling algorithm (voxel neighborhood with 26 adjacency). Finally, only one connected component corresponding to the whole lung was automatically selected (selection corresponding to the second largest component, the first one being the air outside the patient).

This technique does not require any user intervention, except setting the threshold. The measured lung volume is highly dependent on the chosen threshold (e.g., it defines what part of the airways should be included in the lung volume). The choice of an “optimal” threshold was not very relevant to our problem because we only wanted to compare volumes over the three CT images taken from the same patient. Moreover, it is not clear whether such an “optimal” value really exists. Hence, rather than using a single fixed threshold, we chose to compare several lung volume measurements obtained from a range of 10 threshold values (between −200 and −600 Hounsfield units), centered near an automatic threshold, computed by the technique by Hu et al. (38). This produced a total of 8 patients × 3 CT × 10 thresholds = 240 segmentations. The lung volume difference (LVD) between two CT scans of a given patient is given by the following:

\[
LVD(A, B) = \frac{[Vol(A) - Vol(B)] \times 100}{Vol(A)}
\]

where Vol(A) denotes the lung volume of A.

Segmentation variability analysis. Our first study evaluated the influence of threshold variations on the LVD. We first investigated the relation between the threshold and lung volume among the three CT images of each patient. The threshold range was chosen such that the relation was quasi-linear. The slope of this linear
relation and extrema values of lung volumes were computed and compared.

**Lung volume difference analysis.** For each patient, the LVD between each pair of CT scans was computed for each threshold level. The differences are expressed in cubic centimeters and in the percentage of initial lung volume. When the LVD was >7%, we investigated the distribution of this difference between the left and right lungs. Lung segmentation was performed with morphologic mathematics tools. First, the previously computed lung volume was progressively eroded (by a fixed number of erosion steps of a 1-voxel kernel radius) to break the junction between the two lungs. Then, 3D connected components labeling was performed, and the two largest remaining components were labeled as the right and left lung. Finally, the two components were dilated to the initial whole lung volume to propagate the left and right labels. The trachea was also segmented using a similar technique. The mean lung density was computed on each CT scan.

Our objective was not to perform a perfect segmentation but to develop an automated, reproducible technique permitting comparison of volumes using the same basis. Figure 1 shows an example of the segmentation result.

**Motion estimation with nonrigid registration algorithm**

The 3D displacement of each individual lung point between two acquisitions was estimated using image registration methods. Registering two images consists of finding a geometric transformation capable of pairing, as best as possible, each point in one image (called the reference image, denoted by $I$), with its correspondent in the second image (called the object image, denoted by $J$) (39). There are two main classes of transformations (denoted by $T$): rigid (translations, rotations) and nonrigid (allowing local deformations). The complexity of lung motion cannot be reduced to a simple rigid transformation, so we used a 3D nonrigid registration method to estimate the local deformations between CT acquisitions.

Nonrigid methods can be classified into two categories: sparse vector field methods and dense vector field methods. A sparse vector field method requires a set of control points (or landmarks) in the reference image with known correspondence in the object image. The final transformation is computed using a deformation model (such as thin-plate spline [40]) to interpolate point displacements. However, defining or detecting point correspondences between two different acquisitions is difficult in the lung, and the final result depends highly on the accuracy of the corresponding points. Dense vector field methods compute a displacement vector for each point of the volume, leading to potentially more robust point correspondence; no control points were needed.

We used a home-made method based on the algorithm proposed by Thirion (41) and Cachier and Ayache (42, 43). The method (44) is a modified version of the optical flow technique (45) allowing retrieval of large and small displacements. This iterative intensity-based method compares image intensities (gray levels) and requires no segmentation. Each iteration has two steps: pairing and regularization. The corresponding point of $x$ at iteration $i$, denoted by $u_i(x)$, is estimated according to the local gradient of the reference image (in the equation below $J, T$ denotes the composition, $\nabla$ denotes the gradient operator, $\alpha$ is a parameter allowing control of the maximal local displacement in one iteration). This approach relies on the assumption of a gray level intensity conservation between two acquisitions (both images are acquired with the same protocol). The regularization step consists of avoiding spatial incoherences (e.g., two neighboring points with opposite motions). We used a 3D gaussian recursive filter (46) that was shown to be similar (under some assumptions) to the viscous fluid regularization method (47).

$$u_i(x) = \frac{I(x) - J_iT_{i-1}(x)}{\|\nabla I\|^2 + \alpha^2 (I(x) - J_iT_{i-1}(x))^2} \nabla I$$

(2)

This method results in a dense 3D deformation field, computed every 2 mm in the in-plane dimension, between two CT scans: at each voxel of the first CT scan, a 3D vector points to the corresponding voxel in the second CT scan.

**Validation.** Validation of deformable registration is a difficult issue because generally no gold standard is available. Several
approaches exist (48), including definition of a "bronze standard" based on expert analysis, use of synthetic data, and consistency validation. First, we asked a physician to select identifiable points in each volume manually and compare selected points to points displaced with vector field. Such a technique is time-consuming and error prone (49). It was only used for 3 patients and did not lead to any discrepancies. Validation using synthetic data (known warping function applied to a volume) was also attempted and did not show any discrepancies. Finally, we studied the consistency of the resulting transformations with real data. The algorithm was not symmetric: A to B registration did not necessarily lead to the same result as B to A registration. Thus, one way to assess the consistency of the results is to check whether the two deformable transformations \( T_{AB} \) and \( T_{BA} \) are symmetric and whether three transformations \( T_{AB}, T_{BC}, \) and \( T_{CA} \) are transitive \( (T_{AB} + T_{BC} = T_{CA}) \). Preliminary results have shown that the results are consistent (mean symmetric error around 1.3 mm and a mean transitivity error around 1.8 mm). Additional studies are in progress to quantify deformable registration uncertainties.

**Residual error.** The resulting deformation field is composed of two parts: a global rigid deformation (corresponding to the misalignment of the patient between the two acquisitions) and a deformation that we called the residual error (corresponding to lung motion differences between two breath holds). To determine this residual error, we applied a 3D rigid registration algorithm (50, 51) that we adapted to favor the registration of the rigid bony structures. Accuracy was verified by visual checking. Deformable registration was performed using a starting vector field computed from the transformation obtained with rigid registration. The residual error was then obtained by subtracting this rigid transformation from the deformation field obtained by nonrigid registration. It is also possible to resample on image according to the rigid transformation obtained and to start deformable registration from this resampled image. However, it requires an additional interpolation. However, we did not compare the two approaches in detail.

**Data computation.** Each CT acquisition was alternatively used as object or reference image. For each patient, we computed 6 rigid and 6 nonrigid registrations, leading to a total of 48 rigid and 48 nonrigid registrations. Each vector fields represents \( 256 \times 256 \times 70 \) = 4.6 million vectors, amounting to 52.5 Mb in total (each vector is defined by 4 bytes per coordinate). There are from 870,000 to 2 million vectors for each lung, depending on the patient. We also divided the lung volume into six regions (from inferior to superior): inferior 10%, intermediate 20% \( \times 4 \), and superior 10%, as described in Remouchamps et al. (17). We analyzed the displacement by computing the mean norm displacements inside the lung and on the lung surface only. For comparison purposes with their study (17), we developed an algorithm to compute the distance to agreement (DTA). The DTA between two lung surfaces averages the distance of each point on the first lung surface to the closest point on the second surface. Our method used a distance map algorithm (52) to compute the distance of each voxel in a volume to the closest point on the lung surface. The final DTA computation was performed by summing the distances obtained for points belonging to the surface. The computation time for the whole process (segmentation, vector fields, measurements) with a \( 512 \times 512 \times 70 \) image took \(<30 \text{ min using a standard personal computer (2.7 Gz, 1 Go RAM, Linux).}

**Visualization.** Dense 3D vector fields are difficult to visualize because of the amount of information provided. We generated three kinds of images for better visualization and interpretation of the residual error. Figure 2 presents projected 3D vector fields on two slices. Arrows corresponded to the residual error between two breath holds, showing the displacement of the surrounding points. Density windowing of the slices was adapted for visual purpose to observe both low (inside lung) and high (rigid structures) density points. The 2D projection of the vector field fails to show the third dimension of the information: the normal displacement of the projected vector. The axial slice is shown again in Fig. 3, in which two colors depict the cranio-caudal displacement. Green corresponds to displacements toward the patient’s head and blue toward the patient’s foot. Green and blue intensities are scaled as a function of the norm of the displacement Fig. 4 illustrates the norm of the displacement vectors (dark red corresponds to small and light red to large displacements).

**Residual gross tumor volume motion.** Vector fields can be used to quantify motion inside and around the tumor. We modeled the tumor region by an ellipsoid parameterized with center and three axes. The ellipsoid manually selected by the physicians, encompassed the gross tumor volume (GTV; excluding nodes) and was centered at the GTV center. We tried to define the ellipsoid as close as possible to the tumor volume contoured by the physicians. Selection was performed on the reference CT scan that was used for treatment planning. We then considered the residual GTV motion as the set of vectors extracted from the original vector fields belonging to this ellipsoid.

**RESULTS**

**Lung volume**

**Segmentation variability analysis.** The lung volume is quasilinear according to the threshold (mean of asymptotic standard errors \(<0.3\%\)). The relationship between the threshold and volume remained quasiconstant among the three CT scans of each patient (differences between extrema volumes were \(<60 \text{ cm}^3 \) or 4\%), but differed from patient to patient (the slope of the linear relationship between threshold and volume ranged from 1.1 to 2.0).

The volume variations according to the different thresholds were very consistent. For each patient, the standard deviation (for the 10 segmentations) between LVDs (expressed as the percentage of the initial volume) varied from 0.1\% to 1.2\%, suggesting that the LVD remained quasiconstant within the range of thresholds considered.

**LVD analysis.** The LVD ranged from 8 (0.2\%) to \( \approx 1000 \text{ cm}^3 \) (16\%). Table 2 shows the LVD for each patient and each of the three CT scans, in both cubic centimeters and percentages. Table 3 depicts the LVD distribution in the left and right lungs for Patients 3 and 4.

The mean lung density was 0.17 g/cm\(^3\) (SD 0.02) for all images acquired using ABC, and 0.33 g/cm\(^3\) (SD 0.15) for free-breathing scans (Table 4). The increase in lung volume due to the DIBH ranged from 7.7\% to 34.5\%.

**Lung deformation**

Table 5 shows the displacements of all points in the lung. It also indicates the mean of the 5\% points with the greatest displacement. For each patient, the three CT scans are denoted by A, B, and C. Each value in Table 5 was com-
puted according to a mean of the six vector fields (AB, BA, AC, CA, BC, and CB).

Table 6 shows the same information but only for points located inside an ellipsoid covering the GTV. It also gives the tumor size and localization.

Figure 5 represents the mean point displacements for
each patient in the AP, medial–lateral and craniocaudal directions, as well as the 3D norm.

Table 7 represents the mean point displacements in each of the six successive vertical lung regions for the left and right lung (superior 10%, four intermediate 20%, and inferior 10%). For each region, we depicted the displacements computed within the lung volumes, computed on lung surfaces, and computed with DTA.

**DISCUSSION**

Previous studies on breath-hold techniques have shown the necessity of a good reproducibility and the difficulty to evaluate this reproducibility (17, 27, 30, 36). Both two-dimensional and 3D techniques can be used. Two-dimensional techniques involve the use of radiographic films (26, 34) or portal images (31, 53). Evaluation was performed by measuring features such as the projection of the top diaphragm cupola relatively to the skeleton (assumed to be fixed) (21, 26, 29, 34) or implanted radiopaque markers (26). 3D techniques can compare several CT scans acquired at equivalent stages. They were used to compare LVDs (27, 30, 36), lung-surface distances (17, 30) (using the “A not B; B not A” technique or DTA), or feature points, such as the trachea,
carina, tumor center, or diaphragm, that are (mostly manually) localized on each scan (54–56).

The results are expressed in terms of the craniocaudal, AP, or medial–lateral displacements for two-dimensional studies, and lung volume percentage differences or mean 3D displacements for 3D studies. Reproducibility generally varies, from 1.0 (53) to 6.6 mm (26). Table 1 depicts the results of the different studies evaluating

Fig. 3. Craniocaudal (CC) displacement on axial slice. Green corresponds to displacements toward patient’s head and blue toward patient’s foot.

Fig. 4. Norm of 3D vectors on axial slice. Dark red corresponds to high displacement, light red to low norm.
breath-hold interfraction reproducibility, with or without ABC.

**Volume analysis**

Patients were stratified into two groups (paired t test showed that the differences in the two sets were statistically significant); the first one (Patients 1, 2, and 5 to 8) had good lung volume reproducibility compared with the second (Patients 3 and 4).

The first group had results comparable to those of previous studies (LVD < 4.1% or 170 cm³ for the whole lung). Stromberg et al. (36) found 4% mean lung volume differences (at deep inspiration) for intra- and inter-sessions. Wong et al. (30) found approximately 6% intrafraction (30 min apart) variations of lung volume for 3 patients. Wilson et al. (27) found interfraction LVDs varying from 0.2% to 8.7% (< 186 cm³) for the right or left lung (differences not statistically significant, 10 patients).

Greater differences were reported in 2 patients (6–16%). Wilson et al. (27) mentioned 1 patient with a 13.2% (289 cm³) LVD for the left lung (according to the authors, the difference seemed to be due to restitution of lung volume owing to tumor response treatment rather than failure of ABC). The results reported for Patient 3 showed that the difference was almost equally distributed (8.4%) in each lung, and in Patient 4 most of the difference was localized in the left lung (28%).

In 2 patients, the discrepancies observed could have had clinical causes. Patient 3 had a large emphysema bulla of about 512 cm³ in the left lung and numerous other smaller emphysema foci near the apex. Moreover, he had a pleural effusion that increased between acquisitions and right lower lobe atelectasis. Patient 4 also had atelectasis and decreasing pleural effusion, together with a very low forced expirations volume in 1s (0.7). It seems that patients with pleural effusion or atelectasis should not be treated using the ABC device.

In our study, the patient lung volume increased with the breath hold compared with free breathing (average 25%), but remained below the levels reported in the study by Cheung et al. (11) (average 42%, range 23–66%) and comparable to the results presented in by Hanley et al. (54). This lower increase was probably a result of two reasons: all the patients had severe respiratory insufficiency and we used medium DIBH rather than DIBH.

**Lung deformation**

As expected, we first observed a correlation between LVD and mean displacements, with the mean increasing as a function of the volume. However, 2 patients (patients 3 and 5) had similar mean displacements (4.3 mm) but different LVDs (4.8% vs. 8.4%).

We observed fewer displacements in the upper regions of the lung than in the lower parts (Table 7), as was also observed by Remouchamps et al. (17). However, the displacements observed in our study were greater than in their study (17), and the differences between the upper and lower regions were also greater (1.9 to 4 vs. 1.5 to 2.1 mean left to right). Our approach had two major differences with the one by Remouchamps et al. (17). First, we averaged the distance for each point in the volume and not only on the surface. Second, we computed the estimated displacement of each point, not its distance to the closest point. This technique avoids a potential drawback of DTA that tends to underestimate motion (e.g., in the case of vertical sliding). In Table 7, DTA computation showed lower values than the computed displacements. Volume computation also makes it possible to skip the step of internal extrapolation needed in the surface-based approach.

An important point was the elimination of global misalignment between the different CT acquisitions. Most scans had very little differences (<1°, 2 mm), except for some scans acquired at an interval of several days (30 mm).

Residual displacements inside the GTV ellipsoid showed how clinically useful results can be derived from

---

### Table 2. Lung volume differences for three CT scans (A, B, and C) of 8 patients

<table>
<thead>
<tr>
<th>Pt. No.</th>
<th>A–B (cm³)</th>
<th>B–C (cm³)</th>
<th>A–C (cm³)</th>
<th>A–B (%)</th>
<th>B–C (%)</th>
<th>A–C (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>148.6</td>
<td>16.5</td>
<td>165.1</td>
<td>3.9</td>
<td>0.4</td>
<td>4.1</td>
</tr>
<tr>
<td>2</td>
<td>159.5</td>
<td>80.8</td>
<td>78.7</td>
<td>1.8</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>3</td>
<td>343.2</td>
<td>638.8</td>
<td>309.9</td>
<td>6.6</td>
<td>13.1</td>
<td>5.6</td>
</tr>
<tr>
<td>4</td>
<td>603.4</td>
<td>1015.5</td>
<td>412.1</td>
<td>8.9</td>
<td>16.4</td>
<td>5.7</td>
</tr>
<tr>
<td>5</td>
<td>134.3</td>
<td>261.4</td>
<td>395.7</td>
<td>1.6</td>
<td>3.2</td>
<td>4.8</td>
</tr>
<tr>
<td>6</td>
<td>27.1</td>
<td>72.6</td>
<td>96.6</td>
<td>0.5</td>
<td>1.3</td>
<td>1.8</td>
</tr>
<tr>
<td>7</td>
<td>157.9</td>
<td>117.1</td>
<td>40.8</td>
<td>2.8</td>
<td>2.1</td>
<td>0.7</td>
</tr>
<tr>
<td>8</td>
<td>47.1</td>
<td>105.7</td>
<td>58.6</td>
<td>0.7</td>
<td>1.6</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Percentage is percentage of initial volume.

### Table 3. Lung volume difference for left and right lungs in 2 patients with large LVD

<table>
<thead>
<tr>
<th>Pt. No.</th>
<th>Whole lung (cm³)</th>
<th>Right lung (cm³)</th>
<th>Left lung (cm³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>431.2 (8.5)</td>
<td>284.2 (8.5)</td>
<td>146.5 (8.4)</td>
</tr>
<tr>
<td>4</td>
<td>698.2 (10.6)</td>
<td>196.8 (4.2)</td>
<td>506.6 (28.1)</td>
</tr>
</tbody>
</table>

*Abbreviation: LVD = lung volume difference. Data in parentheses are percentages.*
vector fields. The displacements seemed to be slightly lower than those observed inside the whole lungs for Patients 2, 5, 6, 7, 8. However, for Patient 1, taking into account the GTV residual motion, the reproducibility and was good between scans B and C (1.3 \( \pm 0.3 \) mm) but not between B and A (5.9 \( \pm 0.7 \) mm). This case shows that the reproducibility results obtained in the whole lung could not be easily generalized to the tumor itself. We observed no clear correlation between tumor localization in the lung and the GTV residual motion. Yet, work is in progress and detailed dosimetric studies will be presented elsewhere.

The drawbacks of the proposed method are as follows: precision depends on the slice thickness (5 mm here), which may overestimate the displacements in the Cranio-caudal direction. We used linear interpolation throughout the computation step, which tends to smooth high gradients. A smaller slice thickness (3 mm) or high-order interpolation methods (such as cubic spline) can improve accuracy. Another drawback was that we do not know whether the displacements observed arose from breath-hold differences or from anatomic changes between CT scans (e.g., tumor regression). Some studies have evaluated the influence of cardiac motion on the treatment of tumors located near the heart while patient’s lungs were immobilized. The results (1, 5, 9, 57) have shown that heart beats may not have statistically significant impact on tumor motion. We did not investigate this effect.

The method we proposed also presents valuable advantages. First, it is automated. Second, it does not require the determination of corresponding points in each CT scan and, consequently, does not depend on the accuracy of this selection. Additionally, it measures 3D information in the whole lung volume: studying the region around the tumor will help calculate internal margins. By knowing lung displacements, it will become possible to determine internal margins and thus to compare the efficiency of the irradiation procedure during breath hold and free breathing.

**CONCLUSION**

Breath-hold techniques are promising, but reproducibility evaluations are a prerequisite for precisely defining internal margins. In this study, we have proposed an original method for evaluating 3D interfraction breath-hold reproducibility. It relies on both rigid and nonrigid registration methods, allowing computation of the "residual error" corresponding to the 3D displacement of each point of a CT scan. We reported the results based on CT scans from patients enrolled in a Phase II trial, for whom three CT scan images were acquired during breath hold using an ABC device. We also analyzed the LVDs and patient compliance.

Active breath-hold with ABC was generally well tolerated, even in patients with severe respiratory insufficiency. Of the 11 patients with severe respiratory insuf-
ficiency, 1 was excluded because he was unable to undergo the treatment, owing to an inability to understand the ABC procedure and 2 patients had incomplete data sets owing to storage failure. For 6 patients, the breath hold was effective and reproducibility was comparable to that in other interfraction studies. Finally, 2 patients showed insufficient reproducibility, with discrepancies due to clinical reasons. The preliminary results on GTV residual motion have shown that tumor residual motion cannot be easily inferred from its localization.

The 3D automated computation tools presented here, such as lung volume measurement and deformation field computation, allow for personalized interfraction evaluation of breath-hold reproducibility. It should be interesting to extend these tools to the evaluation of intrafraction reproducibility, based, for example, on overlapping slices, such as was done in the work described by Hanley et al. (54). The drawbacks include the great thickness of the slices, leading to potential inaccuracy in the craniocaudal direction, the need to

<table>
<thead>
<tr>
<th>Pt. No.</th>
<th>CT scan*</th>
<th>Mean (mm) (SD)</th>
<th>Maximum of norm (mm)</th>
<th>GTV ellipsoid size (cm³)</th>
<th>Tumor location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>B–A</td>
<td>5.9 (0.7)</td>
<td>8.8</td>
<td>4.8</td>
<td>Lower left lobe</td>
</tr>
<tr>
<td></td>
<td>B–C</td>
<td>1.3 (0.3)</td>
<td>2.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>B–A</td>
<td>2.5 (0.8)</td>
<td>5.5</td>
<td>27.0</td>
<td>Lower paratracheal</td>
</tr>
<tr>
<td></td>
<td>B–C</td>
<td>0.9 (0.4)</td>
<td>2.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>B–A</td>
<td>7.1 (1.6)</td>
<td>10.8</td>
<td>5.3</td>
<td>Lower right lobe</td>
</tr>
<tr>
<td></td>
<td>B–C</td>
<td>7.6 (1.4)</td>
<td>11.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>B–A</td>
<td>5.1 (1.1)</td>
<td>9.5</td>
<td>190.1</td>
<td>Upper right lobe</td>
</tr>
<tr>
<td></td>
<td>B–C</td>
<td>11.8 (4.0)</td>
<td>31.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>C–A</td>
<td>3.9 (1.4)</td>
<td>8.9</td>
<td>87.1</td>
<td>Left hilar</td>
</tr>
<tr>
<td></td>
<td>C–B</td>
<td>2.9 (1.3)</td>
<td>8.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>A–B</td>
<td>2.0 (1.2)</td>
<td>6.3</td>
<td>73.6</td>
<td>Upper right lobe</td>
</tr>
<tr>
<td></td>
<td>A–C</td>
<td>2.5 (0.9)</td>
<td>6.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>B–A</td>
<td>2.0 (1.2)</td>
<td>4.9</td>
<td>22.8</td>
<td>Upper left lobe</td>
</tr>
<tr>
<td></td>
<td>B–C</td>
<td>1.4 (0.4)</td>
<td>2.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>A–B</td>
<td>2.3 (0.7)</td>
<td>5.1</td>
<td>251.3</td>
<td>Median left lobe</td>
</tr>
<tr>
<td></td>
<td>A–C</td>
<td>3.0 (1.3)</td>
<td>8.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: GTV = gross tumor volume.
* Deformation field used (A–B means registration between CT scan A and CT scan B, with A as reference).

Fig. 5. Mean point displacements for each patient in three directions: X, medial–lateral; Y, AP; Z, craniocaudal), and 3D norm.
acquire several CT scans, and the large amount of data to process.

The results of this study have shown the importance of quantifying internal displacements that vary with each patient’s respiratory capacity. Because the tools described provide 3D information on each part of the patient’s body, the technique can be used to adapt internal margins. It can also be used to quantify patient anatomic evolution during treatment, provided that CT scans are acquired regularly (e.g., each week). We also plan to apply this algorithm to CT scans acquired in breath hold at different phases of the breathing cycle (44) to build a patient-adapted 3D breathing model permitting to derive four-dimensional dosimetric studies.

REFERENCES

12. Tsujino K, Hirota S, Endo M, et al. Predictive value of...


