

Detection of patient setup errors with a portal image – DRR registration software application

Kenneth Sutherland,^{1a} Masayori Ishikawa,¹ Gerard Bengua,² Yoichi M. Ito,¹ Yoshiko Miyamoto,¹ and Hiroki Shirato^{1,2}

Hokkaido University Graduate School of Medicine,¹ Sapporo 060-8648, Japan;

Hokkaido University Hospital,² Sapporo 060-8648, Japan

kensuth@med.hokudai.ac.jp

Received 25 October, 2010; accepted 13 February, 2011

The purpose of this study was to evaluate a custom portal image — digitally reconstructed radiograph (DRR) registration software application. The software works by transforming the portal image into the coordinate space of the DRR image using three control points placed on each image by the user, and displaying the fused image. In order to test statistically that the software actually improves setup error estimation, an intra- and interobserver phantom study was performed. Portal images of anthropomorphic thoracic and pelvis phantoms with virtually placed irradiation fields at known setup errors were prepared. A group of five doctors was first asked to estimate the setup errors by examining the portal and DRR image side-by-side, not using the software. A second group of four technicians then estimated the same set of images using the registration software. These two groups of human subjects were then compared with an auto-registration feature of the software, which is based on the mutual information between the portal and DRR images. For the thoracic case, the average distance between the actual setup error and the estimated error was 4.3 ± 3.0 mm for doctors using the side-by-side method, 2.1 ± 2.4 mm for technicians using the registration method, and 0.8 ± 0.4 mm for the automatic algorithm. For the pelvis case, the average distance between the actual setup error and estimated error was 2.0 ± 0.5 mm for the doctors using the side-by-side method, 2.5 ± 0.4 mm for technicians using the registration method, and 2.0 ± 1.0 mm for the automatic algorithm. The ability of humans to estimate offset values improved statistically using our software for the chest phantom that we tested. Setup error estimation was further improved using our automatic error estimation algorithm. Estimations were not statistically different for the pelvis case. Consistency improved using the software for both the chest and pelvis phantoms. We also tested the automatic algorithm with a database of over 5,000 clinical cases from our hospital. The algorithm performed well for head and breast but performed poorly for pelvis cases, probably due to lack of contrast in the megavoltage portal image. The software incorporates an original algorithm to fuse portal and DRR images, which we describe in detail. The offset optimization algorithm used in the automatic mode of operation is also unique, and may be useful if the contrast of the portal images can be improved.

PACS numbers: 87.55.Qr, 87.57.nj

Key words: patient setup, image registration, portal imaging, mutual information

^a Corresponding author: Kenneth Sutherland, Hokkaido University Graduate School of Medicine, Sapporo-shi Kita-ku Kita 14 Jo Nishi 5 Chome, 060-8648 Japan; phone:+81-11-706-7638; fax:+81-11-706-7639; email: kensuth@med.hokudai.ac.jp

I. INTRODUCTION

Traditionally doctors have judged patient setup errors by viewing portal images alongside planning digital reconstructed radiograph (DRR) images, either with paper printouts, films on a light board, or on a computer terminal. Estimation of the error is made by measuring the distance from the isocenter to anatomical structures (usually bones) visible in both images.⁽¹⁾ However, the process is often inaccurate, with errors between 5 and 10 mm being reported.⁽²⁾

Fully and semi-automatic methods based on electronic portal imaging devices (EPIDs),⁽³⁾ implanted fiducial surrogate markers imaged with kilovoltage fluoroscopy,⁽⁴⁾ on-board imager (OBI, Varian Medical Systems, Palo Alto, CA)⁽⁵⁾ and cone-beam CT (CBCT), have recently become common. However, these methods often require additional cost, exposure of X-rays, and longer time for setup. As a result, for a large number of patients who do not require a high accuracy of patient positioning (e.g., palliative treatments or mantle field radiation), portal imaging once or twice during the first week of treatment is still desirable.

The standard patient treatment regime employed at our institution involves first obtaining a CT scan of the patient which will be used for treatment planning. A DRR reconstruction of the beam's eye view (BEV) is computed and stored by a commercial treatment planning system (TPS). The treatment port, isocenter, orthogonal axes with scale information (a tick mark each centimeter) and other treatment information are burned into the DRR image.

Portal images are usually obtained at our hospital before the first treatment fraction using a megavoltage X-ray source from a linear accelerator and a computed radiograph (CR) system (Fuji Medical Co., Ltd., Tokyo, Japan). Images are captured onto photosensitive plates, which are scanned to produce high resolution (usually 1760×1760 pixels) deep bit (10 bits per pixel) images. The portal images also contain scale information on the orthogonal axes with a small dot each centimeter. The portal and DRR images are compared to determine that the treatment beam accurately targets the planned treatment volume (PTV) while avoiding organs at risk (OARs). If a problem is detected, the presiding physician may request that the couch position be adjusted. Portal images are then retaken and rechecked against the DRR.

We developed an image registration software application for the estimation of patient setup error. DRRs from any commercial TPS can be opened using common file formats (e.g., bitmap, JPEG, DICOM). The software works for any anatomical region or gantry angle. The software can be operated manually, or with an automatic registration mode based on the mutual information between the images.

The purpose of this study is to verify that our software actually improves setup error detection compared with the traditional side-by-side method. There is little statistical evidence in the literature of the superiority of image registration to side-by-side human estimation of setup error from two-dimensional portal images. We evaluated the efficacy of the software as an aid for the clinical staff to improve setup accuracy using a prospective phantom study and its statistical analysis. We compared the ability of humans to correctly determine a known setup error with and without the software. An automatic mode of operation of the software was also tested with a database of clinical cases, collected over several years, for which the setup up error was determined by the consensus of a trained software operator (medical physicist), a radiation technician and the presiding oncologist.

II. MATERIALS AND METHODS

A. Patient setup error detection software

After a CT scan of the patient is obtained, a DRR of the BEV is computed and stored with a TPS. The treatment port, isocenter, orthogonal axes with scale information (a tick mark each centimeter) and other treatment information are burned into the image. The DRR image is

saved in a database accessible to our software. Portal images are usually obtained before the first treatment fraction using a double exposure (open field and the actual field) with a megavoltage X-ray source from a linear accelerator, captured with a CR system. The portal images also contain graphical scale information on the orthogonal axes with a small dot or tick mark every centimeter from the isocenter.

In our proposed method, while the patient is lying on the treatment couch, the planning DRR and portal images are loaded into the software. The isocenter and two points on orthogonal axes, usually 10 cm from the isocenter, are designated by clicking on the images with the mouse. The locations of the points are determined by the operator using the axis and scale tick marks on each image. The software uses the three control points to determine scaling and rotation in order to transform the portal image into the coordinate space of the DRR image. Because the imaging plate may not be exactly orthogonal to the beam axis, especially when an oblique gantry angle is used, the portal image may be warped. The software can correct for these out-of-plane rotations.

After the images have been successfully fused, the operator uses bony landmarks or other visible anatomical features in order to determine setup error. The portal image can be shifted horizontally or vertically and rotated clockwise or counterclockwise relative to the stationary DRR image. An example screenshot is shown in Fig. 1.

The operator can apply complementary colors, such as red and cyan, to make the fused image more distinguishable. If the images align exactly, each resulting pixel will be a level of grey. Misalignments are visible as color shadows. The operator can specify which pair of complementary colors to apply, depending on which colors are easier to see. Brightness and contrast can also be adjusted on both images to make bony features more distinguishable. The user can also rapidly flip between the DRR, portal and fused images. After the images are successfully registered, the treatment field shift (in mm) and rotation (in degrees), corresponding to the BEV, is displayed. The required couch movement to correct the offset, based on the gantry angle, can also be computed and displayed.

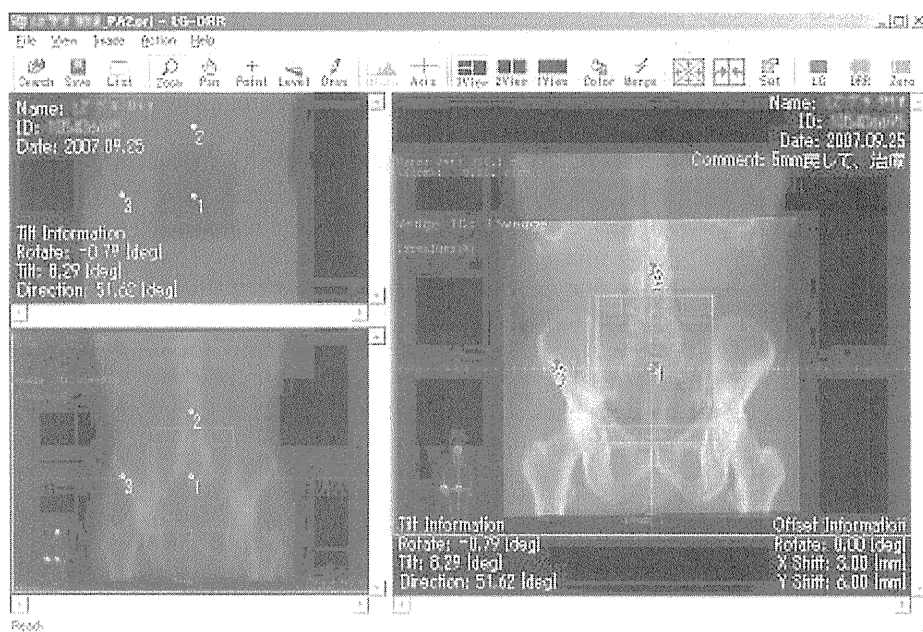


FIG. 1. The Portal-DRR software. The user specifies three points on the portal (top left) and DRR (bottom left) images using grid spacing tick marks. The merged image (right) is used to compare internal bony structures in order to determine the setup error, which is displayed at the lower right.

B. Theory of image distortion correction

Distortions of portal images may include rotations about the beam axis (z-axis) and rotations about an axis perpendicular to the beam axis (x- or y-axis). It is difficult to independently correct rotations about the x and y axes. We therefore devised a scheme to treat rotations about the x and y axes as equivalent rotations about an arbitrary axis perpendicular to the z-axis, denoted as the x' -axis. The x' -axis lies in the x, y plane and points in the direction of the tilt. It is possible to compose rotations about the x' -axis from equivalent rotations about the x and y axes, combined with a rotation about the z-axis. Coordinates on the imaging plate (x' , y') correspond to the original coordinates (x, y) as follows:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = b \cdot Rot(\theta) \cdot Rot(\varphi) \cdot St(\phi) \cdot Rot^{-1}(\varphi) \begin{pmatrix} x \\ y \end{pmatrix} \quad (1)$$

Here b is the zoom ratio, $Rot(\theta)$ is the rotation about the z-axis, $Rot(\phi)$ is the rotation about the x' -axis, and $St(\Theta)$ is the stretch ratio produced by the rotation about the x-axis. The rotation and stretch functions are defined as:

$$Rot(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, St(\phi) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\cos \phi} \end{pmatrix} \quad (2)$$

One can obtain the original image pixel coordinates by applying the reverse process in order to correct the distorted image:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{b} \cdot Rot(\varphi) \cdot St^{-1}(\phi) \cdot Rot^{-1}(\varphi) \cdot Rot^{-1}(\theta) \begin{pmatrix} x' \\ y' \end{pmatrix} \quad (3)$$

Because parameters θ , ϕ , and Θ can be obtained from two arbitrary coordinates, according to the equations, we can proceed as follows.

$$\begin{aligned} \varphi &= \arctan \left(\frac{1}{2} \left(\alpha + \sqrt{4 + \alpha^2} \right) \right) & \varphi + \theta &= \arctan \left(\frac{m \cos \varphi + n \sin \varphi}{k \cos \varphi + l \sin \varphi} \right) \\ \phi &= \cos^{-1} \left(\frac{m \cos \varphi + n \sin \varphi}{k \sin \varphi - l \cos \varphi} \right) & b &= \frac{\cos(\varphi + \theta)}{k \cos \varphi + l \sin \varphi} \end{aligned} \quad (4)$$

Note that k , l , m and n are in the original beam axis' coordinate system and can be computed from two pairs of points.

$$\begin{pmatrix} k & l \\ m & n \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \begin{pmatrix} x'_1 & x'_2 \\ y'_1 & y'_2 \end{pmatrix}^{-1} = \frac{1}{x'_1 y'_2 - x'_2 y'_1} \begin{pmatrix} x_1 x'_1 + x_2 y'_1 & x_1 x'_2 + x_2 y'_2 \\ y_1 x'_1 + y_2 y'_1 & y_1 x'_2 + y_2 y'_2 \end{pmatrix} \quad (5)$$

The conversion process is illustrated in Fig. 2.

The proposed image-registration method uses imaging plates to obtain portal images for irradiation field verification. For non-zero gantry angles, the imaging plate is held freely on a stand without fixation to a certain coordinate system. In order to correct perspective distortion due to the imaging plate not being orthogonal to the beam axis, three control points are

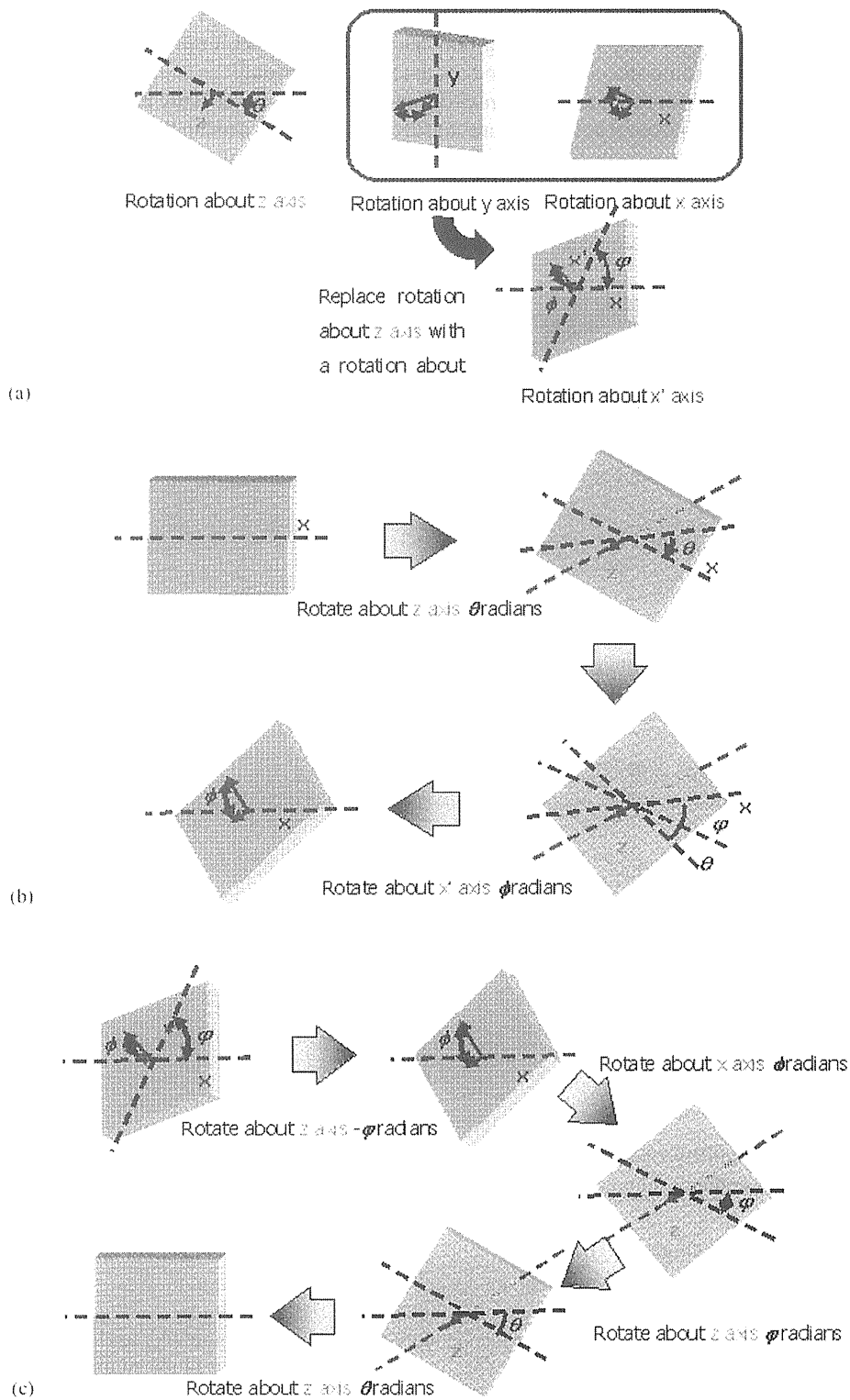


Fig. 2. Conversion of rotation axis for image distortion.

manually placed using graphical axis scale tick marks on each image. The specified points should cover the area of the image containing bony structures used for alignment. We refer to the area roughly surrounding the control points as the “working area”.

To test the correction function, an electronic tiltmeter was used to measure a 10° tilt of the imaging plate about the horizontal axis. The plate was exposed to X-rays using the linac axis grid. The resulting image was fused with a dummy DRR image. A test point was placed at a position 5 cm along the x- and y-axis within the working area. The distance from the test point on the DRR image and the transformed portal image point was measured. The experiment was repeated with the test point outside the working area. The experimental setup is illustrated in Fig. 3.

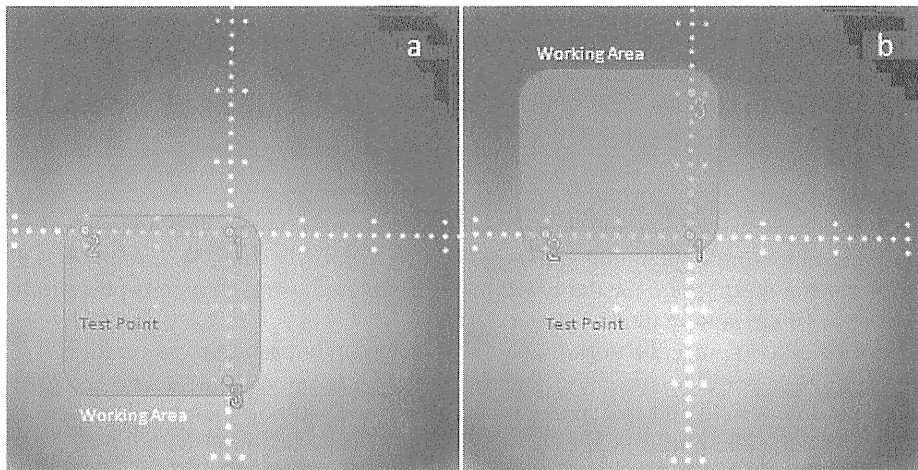


FIG. 3. Tilt experiment setup. The imaging plate was tilted 10° horizontally. The test point (5 cm along the x- and y-axis) was inside the working area for (a) and outside the working area for (b). The test point was transformed into the DRR image coordinates and the distance from the corresponding point specified on the DRR image was measured.

C. Automatic determination of patient setup error

We developed a method to automatically determine the x-y shift and in-plane rotation based on the mutual information (MI) of the pixel intensity values of the overlapping portions of the transformed portal and DRR images. The entropy of a single image can be computed from the equation:

$$H = \sum_i p_i \cdot \log \frac{1}{p_i} \quad (6)$$

where p_i is the probability of a gray-level pixel value i estimated from the histogram of the image.⁽⁶⁾ The joint entropy between two images can likewise be defined as:

$$H(A, B) = - \sum_{i,j} p(i, j) \cdot \log[p(i, j)] \quad (7)$$

where $p(ij)$ is the probability of gray-level pixel value i from one image (A) and the gray-level pixel value j from the second image (B) at the same position. Mutual information is then defined as:

$$I(A, B) = H(A) + H(B) - H(A, B) \quad (8)$$

where $H(A)$ is the entropy of the first image (i.e., the portal image), $H(B)$ is the entropy of the second image (i.e., the DRR image), and $H(A,B)$ is the entropy of their joined histogram.⁽⁷⁾

Optimizing the MI involves finding the x-y shift position that yields the highest mutual information coefficient, automatically shifting the portal image left, right, up and down a determined step size, and choosing the position with the highest MI coefficient. If the current x-y shift value yields the highest MI coefficient, the optimization loop is abandoned. In order to reduce the likelihood of stopping at a local maximum, we employ a multiresolution method. The initial step size is relatively large, usually 4 mm. The best MI coefficient is located with the initial step size. The step size is then reduced by half and the process is repeated. The final step size is usually 1 mm, but can be set lower for submillimeter accuracy.

Adding rotation checks can increase optimization time exponentially if every possible rotation is checked at each shift location. However, if we assume that the setup rotation is small, a good compromise approach is to search rotation and shift iteratively; that is, search for the best x-y shift and then search for the best rotation. This process is repeated until the best possible rotation is found at the best possible x-y shift.

D. Phantom study

CT scans of anthropomorphic thorax and pelvis phantoms were taken. The images were read into a commercial treatment planning software system (Xio; CMS, Inc., St. Louis, MO). DRR images (one for the thorax, one for the pelvis) were generated using an antero-posterior (AP) port with a 10×10 cm square field.

In order to be able to present a portal image with a known but arbitrary x-y error offset, we first obtained a single portal image of the phantom without field or scale ticks. Our test software burns in the treatment field and tick marks based on determined horizontal and vertical offset error values (Fig. 4). The portal images of the phantoms were acquired with a 6 MV photon beam from a linear accelerator. A list of five cases was prepared with offset values between plus and minus 10 mm horizontally and vertically. Rotation was not considered in this study. In order to determine each subject's consistency in determining the offset error, each case was presented twice for a total of ten cases. Each subject was presented with the same offset cases, but the order of the case list was shuffled.

For the ten thorax and ten pelvis cases generated by the computer with the above method, we presented the series of cases to each examiner using three methods. First, five licensed radiation oncologists determined setup error with the conventional side-by-side method, without the aid of the registration software. We refer to this as the "side-by-side method". These doctors included experienced physicians, as well as doctors in training. Four radiation technicians were

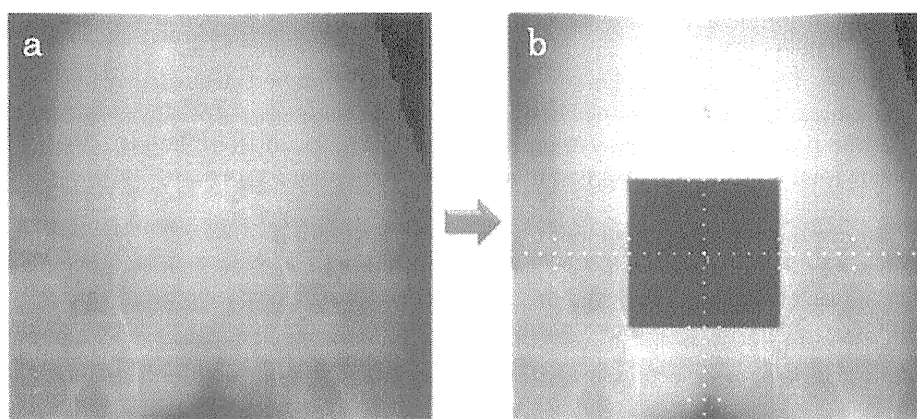


FIG. 4. Portal image taken without field or scale tick marks (a). Automatically generated treatment field and scale ticks applied to the phantom portal image (b).

then asked to determine the same series of setup errors using the registration software, which we refer to as the “registration method”. The automatic determination function of the software was also used to determine the setup error of the same case list (“automatic method”). Doctors and technicians were allowed to adjust image brightness and contrast. No time constraint was in place during the testing.

E. Statistical analysis

The distance from actual and estimated offset was computed with $\sqrt{(x - x_0)^2 + (y - y_0)^2}$ where (x, y) is the estimated offset and (x_0, y_0) is the actual offset. The average and standard deviation of the distance was computed for each examiner.

The average consistency of an examiner was defined as the average geometric distance between two estimations of a single case. The mean of the methods was compared with the paired t-test. The difference of the average consistency among the three methods was also computed in order to determine if consistency improved using the software. All statistical analysis was performed with JMP version 8 (SAS Institute Inc., Cary, NC).

F. Database study

In order to test the performance of the automatic registration mode of the software on actual clinical data, we assembled a database of over 5,000 patient setup cases performed between April 2007 and December 2009 at our hospital. At the time of treatment, the portal image for each case was registered with the corresponding DRR image using the manual mode of the software, and the setup error for each case was determined by a consensus of three people: a software operator (medical physicist), a radiation technician, and the attending oncologist. Although it is difficult to establish a “gold standard” for clinical data, for this study we assume the human-determined offset is correct, or at least very near the actual offset error.

After assembling the database, a large batch script was executed to open each case, fuse the portal and DRR images, and automatically determine the error offset using the MI optimization algorithm described above. The automatically-determined error offset was then compared with the human-determined error offset. The geometric distance between the two values was computed and a statistical analysis was performed.

III. RESULTS

A. Phantom study

The results of the thorax and pelvis are shown in Fig. 5(a) and 5(b), respectively. The geometric distance from the actual offset and the determined offset is plotted. Results from five doctors (DR1–DR5) using the side-by-side method, four technicians (TECH1–TECH4) using the registration software in manual mode, and the automatic method (COMP) are shown.

For the thorax, the average distance between the actual setup error and the estimated error was 4.3 ± 3.0 mm for the radiation oncologists without the registration software, 2.1 ± 2.4 mm for technicians with the registration software, and 0.8 ± 0.4 mm for the automatic algorithm. For the pelvis, the average distance between the actual setup error and estimated error was 2.0 ± 0.5 mm for the radiation oncologists without the registration software, 2.5 ± 0.4 mm for technicians with the registration software, and 2.0 ± 1.0 mm for the automatic algorithm.

Each case was presented to the examiner twice. These two values are plotted vertically for each examiner. Consistency was measured as the geometric distance between the two estimated offsets of the same portal image. This corresponds to the height of the vertical bar in Fig. 5. The results are summarized in Table I.

In order to determine statistically whether setup error estimation improved using the software, we computed the average consistency among the three methods. The average consistency for the side-by-side method was 2.4 ± 2.0 mm for thorax and 1.4 ± 1.2 mm for pelvis. The average

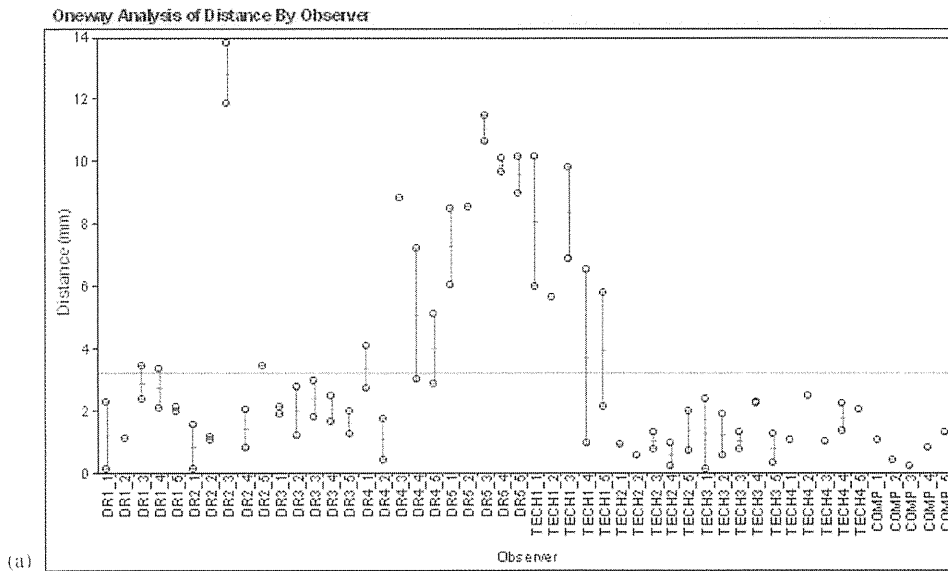


Fig. 5(a). Results of AP chest study.

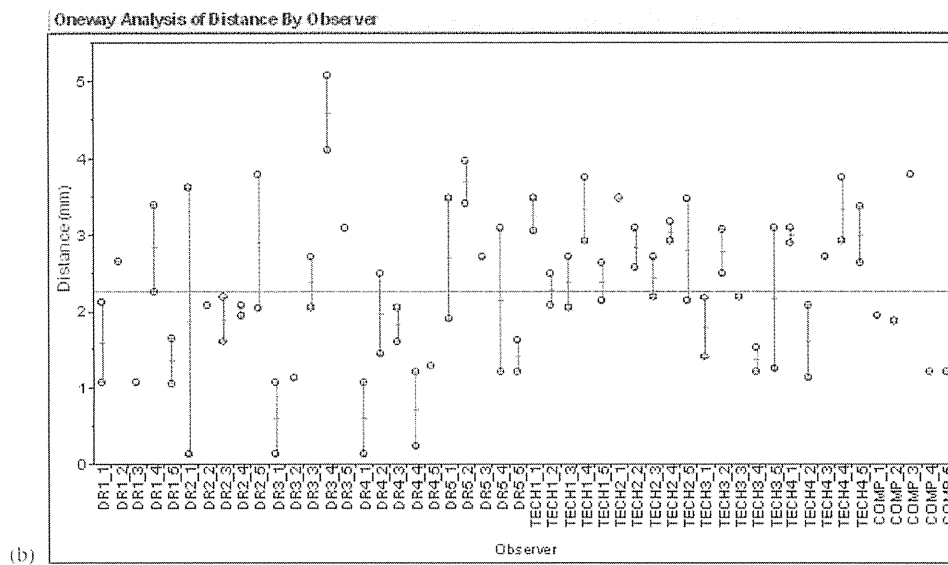


Fig. 5(b). Results of AP pelvis study.

consistency for the registration method was 1.7 ± 1.6 mm for thorax and 0.9 ± 0.5 mm for pelvis. The results are summarized in Table 2. Note that the computer algorithm always produces the same result for a given input, so it is perfectly consistent.

In order to test statistically that setup error estimation improved using the software, we computed the paired two samples for means. The null hypothesis is that population mean of the differences between the paired values is zero. The results are summarized in Tables 3(a) and 3(b).

TABLE 1. Average distance from estimated and actual setup error.

<i>Phantom</i>	<i>Method</i>	<i>Average Distance</i>
Chest	Side-by-side	4.3±3.0 mm
	Registration	2.1±2.4 mm
	Auto	0.8±0.4 mm
Pelvis	Side-by-side	2.0±0.5 mm
	Registration	2.5±0.4 mm
	Auto	2.0±1.0 mm

TABLE 2. Average consistency of test subjects.

<i>Phantom</i>	<i>Method</i>	<i>Average Consistency</i>
Chest	Side-by-side	2.4±2.0 mm
	Registration	1.7±1.6 mm
Pelvis	Side-by-side	1.4±1.2 mm
	Registration	0.9±0.5 mm

TABLE 3(a). Comparison of estimation methods (chest).

<i>Comparison</i>	<i>p-value</i>
Side-by-side - Registration	0.0067
Side-by-side - Auto	0.0002
Registration - Auto	0.0001

TABLE 3b. Comparison of estimation methods (pelvis).

<i>Comparison</i>	<i>p-value</i>
Side-by-side - Registration	0.0047
Side-by-side - Auto	0.4547
Registration - Auto	0.0593

B. Database study

Results of the database study are summarized in Table 4. Overall, the automatic registration method does not perform well at around 7 mm from the human-estimated offset. However, for certain anatomical regions, such as head and breast, the algorithm consistently estimated the setup error within about 2 mm of the human estimation. Automatic registration was especially poor for pelvic regions — more than 1 cm on average. This is probably due to the lack of contrast in the portal images.

In order to demonstrate the usefulness of the automatic mode of operation for two sample anatomical cases, we arbitrarily divided the distance between the human and computer estimated offsets into four bins: ≤ 2 mm (Good), ≤ 5 mm (Fair), ≤ 10 mm (Poor) and > 10 mm (Terrible). Tables 5(a) and 5(b) show the results for head and pelvis cases for all gantry angles.

TABLE 4. Comparison of auto-shift function with clinical case database.

<i>Group</i>	<i>Count</i>	<i>Mean Dist. (mm)</i>
All	5101	7.0±2.1
AP Head	407	2.4±2.1
LR Head	537	2.2±2.7
AP Neck	158	5.5±6.1
LR Neck	214	4.2±4.3
OB Neck	95	4.8±4.1
AP Chest	264	5.1±4.5
LR Chest	101	7.7±8.0
OB Breast	160	3.6±2.7
AP Pelvis	148	13.8±19.8
LR Pelvis	148	12.8±11.7

TABLE 5(a). Comparison of auto-shift function with humans (head).

<i>Bin</i>	<i>Frequency (total = 61)</i>
Good (≤ 2 mm)	35 (57%)
Fair (≤ 5 mm)	23 (38%)
Poor (≤ 10 mm)	1 (2%)
Terrible (> 10 mm)	2 (3%)

TABLE 5(b). Comparison of auto-shift function with humans (pelvis).

<i>Bin</i>	<i>Frequency (total = 48)</i>
Good (≤ 2 mm)	4 (8%)
Fair (≤ 5 mm)	9 (19%)
Poor (≤ 10 mm)	16 (33%)
Terrible (> 10 mm)	19 (40%)

C. Tilt experiment

Results of the tilt experiment are presented in Table 6. When the test point was placed 5 cm within the working area, the distance from the DRR point and the transformed portal image point was 0.7 mm. When the point was placed 5 cm outside the working area, the distance increased to 1.5 mm. This result indicates that the user must be careful to consider only bony structures within the working area when the out-of-plane tilt is large. The automatic registration algorithm should also ignore pixels outside the working area.

TABLE 6. Result of tilt experiment.

<i>Point Placement</i>	<i>Distance</i>
Inside working area	0.7 mm
Outside working area	1.5 mm

IV. DISCUSSION

Our results show that the registration method was at least not worse than the side-by-side method, and the automatic method was statistically better than both the side-by-side method and registration method for the thorax phantom case studied. This result suggests that our

software can be a reasonable complementary method in the clinical practice. Based on this study, the software has been installed in our hospital information system and has been used in clinical practice since April 2007. Moreover, the time required for the estimation of the setup error has not been extended by the usage of this program, once the operator obtained sufficient experience with the software.

A number of criteria influenced our software design decisions:

- Because image registration is performed while the patient is waiting in the treatment position, the system must be fast — less than about one minute after the portal image is obtained.
- DRRs by any commercial planning software are to be used, as long as scale information is burned into the bitmap. The DRR image can even be a screen-captured bitmap, or a digital scan of a paper, or film print.
- Any common file format (e.g., bitmap, JPEG, DICOM) can be used for either the portal or DRR image.
- The system must work with any beam view of any anatomical region encountered in clinical practice.
- The software must run on a single standard PC running the Windows operating system.

We have found that there was a large inconsistency in the side-by-side method among doctors. This is probably due to a difference in training or experience. The large variation in the accuracy of the final decision based on the portal film may influence the clinical outcome. The improvement in the consistency with automatic registration (which was found in this study for thorax) may improve the local control rate and complication rate in this context.

If Dr. 5, in particular, is omitted, the side-by-side method compares well with the registration methods. Although Dr. 5 was a licensed radiation oncologist with experience estimating setup errors, he or she may have needed more practice with our experimental setup. Although we explained that the movement of the radiation field (rather than the treatment couch) should be specified, it is easy to mistake left-right or up-down shift. We feel that these kinds of human errors are inevitable when relying solely on human judgment with the side-by-side method.

From the database study, we can state that the automatic mode of operation performs well for head, neck and breast cases, but performs poorly for pelvic cases. Due to the thickness of the human anatomy in the pelvic and abdomen regions, the contrast of the resulting megavoltage portal image is very low. This makes automatic registration based on mutual information extremely difficult. In order to improve the performance of the automatic registration, the contrast of the portal image needs to be improved. Possible methods of improving portal image contrast include using kilovoltage X-ray, or using advanced digital image processing.⁽⁸⁾ Other methods, such as restricting the area used to compute the mutual information may further improve auto-registration.

Although the automatic registration algorithm performed reasonably well on the pelvis phantom, the brightness and contrast (window level) of the portal image from the phantom image was adjusted by hand to maximize the contrast of the pelvic bone. In the database study, however, it was not feasible to adjust all cases by hand. The window level was computed automatically from the raw scanned data based on the histogram of pixel values. If the window level algorithm can be improved, the image contrast may be increased. Sophisticated image filters may also improve the signal-to-noise ratio and make bony features more recognizable.⁽⁹⁾ We hope that this will, in turn, improve the auto-registration results.

While some verification systems are applicable to only a particular application (for example, lung cancer⁽¹⁰⁾ or pelvic treatments⁽¹¹⁾), our system is used for all cases encountered in our practice, including oblique views. Some systems require the computation of custom DRR images in order to match as closely as possible the portal image.⁽¹²⁾ Our system uses the DRRs that are generated by a commercial TPS software package. Thus, the same pair of portal and DRR images that the attending physicians have normally been using are used by our software.

Some systems require the user to draw bony structures on the source images.⁽¹³⁾ While this is an option with our software, it is not required. It is sometimes useful to draw the outline of a bone on both images, and then see how closely they align on the merged image; however, we found that this method significantly increases the burden on the operator. It is generally faster to view the merged portal-DRR image without hand-drawn contours.

Although EPID has become popular, some EPID systems are quite expensive to maintain and have reduced imaging quality. Our software can be used in many situations where good quality EPID is not available.

Because our system uses only three points to calculate the out-of-plane transformation, areas far from the control points (outside the working area) may not align exactly. If the entire image is to be used for alignment, an alternative method would be to use a calibration function such as *McalList* from the *Matrox Imaging Library*.⁽¹⁴⁾ With this method, a list of corresponding points on the portal and DRR images is passed to the function and a perspective distortion correction matrix is calculated. The more points that are specified, the more accurate the mapping. However, because only points on the axes can be specified, the corners of the corrected portal image may still be slightly distorted. Specifying many points may also be a burden on the user. An automatic method of detecting all the visible grid tick marks on the portal and DRR images is under investigation.

V. CONCLUSIONS

The ability to estimate offset values improved using our software for the chest phantom that we tested. Setup error estimation was further improved using our automatic error estimation algorithm. Estimations were not statistically different for the pelvis case. Comparing the automatic setup function with a database of clinical cases estimated by human operators revealed that the automatic function works relatively well for head, chest and breast cases, but performs poorly for pelvis and other cases. Automatic registration should improve by increasing the contrast of the portal image. Although setup error can be manually judged accurately and quickly with the software as an aid to doctors and technicians, work remains to make the software more fully automatic.

REFERENCES

1. Hurkmans CW, Remeijer P, Lebesque JV, Mijnheer BJ. Set-up verification using portal imaging: review of current clinical practice. *Radiother Oncol.* 2001;58(2):105–20.
2. Clippe S, Sarrut D, Malet C, Miquet S, Ginestet C, Carrie C. Patient setup error measurement using 3D intensity-based image registration techniques. *Int J Radiat Oncol Biol Phys.* 2003;56(1):259–65.
3. Boyer AL, Antonuk L, Fenster A, et al. A review of electronic portal imaging devices (EPIDs). *Med Phys.* 1992;19(1):1–16.
4. Shirato H, Shirnizu S, Kunieda T, et al. Physical aspects of a real-time tumor-tracking system for gated radiotherapy. *Int J Radiat Oncol Biol Phys.* 2000;48(4):1187–95.
5. Sorcini B and Tilikidis A. Clinical application of image-guided radiotherapy, IGRT (on the Varian OBI platform). *Cancer/Radiother.* 2006;10(5):252–57.
6. Pluim J, Maintz J, Viergever M. Mutual-information-based registration of medical images: a survey. *IEEE Trans Med Imaging.* 2003;22(8):986–1004.
7. Tsao J. Interpolation artifacts in multimodality image registration based on maximization of mutual information. *IEEE Trans Med Imaging.* 2003;22(7):854–64.
8. Nakashima K, Ashizawa K, Ochi M, et al. Interpretation of normal anatomical structures on chest radiography: comparison of Fuji computed radiography (FCR) 5501D with FCR 5000 and screen-film system. *J Appl Clin Med Phys.* 2003;4(1):85–90.
9. Yamada S, and Murase K. Effectiveness of flexible noise control image processing for digital portal images using computed radiography. *Br J Radiol.* 2005;78(930):519–27.
10. Van Sörensen de Koste JR, De Boer HCJ, Schuchhard-Schipper RH, Senan S, Hejimen BJ. Procedures for high precision setup verification and correction of lung cancer patients using CT-simulation and digitally reconstructed radiographs (DRR). *Int J Radiat Oncol Biol Phys.* 2003;55(3):804–10.

11. Hanley J, Mageras GS, Sun J, Kutcher GJ. The effects of out-of-plane rotations on two dimensional portal image registration in conformal radiotherapy of the prostate. *Int J Radiat Oncol Biol Phys.* 1995;33(5):1331–43.
12. Dong L, and Boyer A. An image correlation procedure for digitally reconstructed radiographs and electronic portal images. *Int J Radiat Oncol Biol Phys.* 1995;33(5):1053–60.
13. Lujan AE, Balter JM, Ten Haken RK. Determination of rotations in three dimensions using two-dimensional portal image registration. *Med Phys.* 1998;25(5):703–08.
14. Matrox Electronic Systems. Matrox imaging library, version 9: user guide. Manual no. Y10513-301-0900. Dorval (HQ), Canada: Matrox Electronic Systems Ltd.; 2008. p. 200–01.

IDH1/2 gene status defines the prognosis and molecular profiles in patients with grade III gliomas

Ichiyo Shibahara · Yukihiro Sonoda · Masayuki Kanamori · Ryuta Saito · Yoji Yamashita · Toshihiro Kumabe · Mika Watanabe · Hiroyoshi Suzuki · Shunsuke Kato · Chikashi Ishioka · Teiji Tominaga

Received: 13 July 2011 / Accepted: 6 September 2011
© Japan Society of Clinical Oncology 2011

Abstract

Background The discovery of isocitrate dehydrogenase 1 and 2 gene (*IDH1/2*) mutations has enabled grade III glioma to be divided into mutated and wild-type *IDH1/2* groups, which are known to carry different prognosis and molecular features. However, detailed subgroup analysis of grade III glioma is limited. To address this, we investigated molecular and prognostic features of grade III glioma with and without *IDH1/2* mutation.

Methods We retrospectively analyzed 115 grade III glioma patients. Clinical parameters were obtained from medical records. The mutation of *IDH1/2* and *TP53* was analyzed by direct sequencing. O⁶-methylguanine methyltransferase gene (*MGMT*) gene promoter methylation status was determined by methylation-specific polymerase

chain reaction. Detection of chromosome copy number changes of 1p, 7p (*EGFR*), 9p (*CDKN2A*), 10q (*PTEN*), and 19q was carried out by multiple ligation-dependent probe amplification. Patients were divided into two groups, mutated *IDH1/2* and wild-type *IDH1/2*, for correlation with the factors analyzed.

Results In our series, as previously reported, *IDH1/2* mutation was an independent prognostic marker for improved progression-free and overall survival (OS) ($P < 0.0001$ and $P < 0.0001$, respectively) in patients with grade III gliomas. Subgroup analysis found that incomplete resection, 7p gain, and *TP53* mutation were independent prognostic factors of poor outcome in grade III glioma patients with mutated *IDH1/2* ($P = 0.0092$, $P = 0.015$ and $P = 0.026$, respectively), while there were none in patients with wild-type *IDH1/2*.

Conclusions *IDH1/2* gene status was significantly associated with prognosis in grade III gliomas. Subgroup analysis found that poor prognostic factors existed even in patients with *IDH1/2* mutation.

Electronic supplementary material The online version of this article (doi:10.1007/s10147-011-0323-2) contains supplementary material, which is available to authorized users.

I. Shibahara · Y. Sonoda (✉) · M. Kanamori · R. Saito · Y. Yamashita · T. Kumabe · T. Tominaga
Department of Neurosurgery, Tohoku University
School of Medicine, 1-1 Seiryomachi, Aoba-ku,
Sendai 980-8575, Japan
e-mail: sono@nsg.med.tohoku.ac.jp

M. Watanabe
Pathological Division, Tohoku University Hospital,
1-1 Seiryomachi, Aoba-ku, Sendai 980-8575, Japan

H. Suzuki
Department of Pathology, Sendai Medical Center,
2-8-8 Miyagino, Miyagino-ku, Sendai 983-8520, Japan

S. Kato · C. Ishioka
Department of Clinical Oncology, Tohoku University
School of Medicine, 1-1 Seiryomachi, Aoba-ku,
Sendai 980-8575, Japan

Keywords Grade III gliomas · *IDH1/2* · 1p/19q · *TP53* · 7p gain

Introduction

Recently, total genome analysis of glioblastoma (grade IV) identified somatic mutations at codon 132 of the isocitrate dehydrogenase 1 gene (*IDH1*) and codon 172 of the isocitrate dehydrogenase 2 gene (*IDH2*) [1–3]. These gene mutations occur at a very early stage of gliomagenesis [1, 4], and therefore the status of the *IDH1/2* gene can be used to classify glioma into two distinct types, mutated *IDH1/2* and wild-type *IDH1/2*, which have implications for

the scheme of the glioma development [5, 6]. Currently, these two distinct types carry significantly different molecular characteristics, as Yan et al. [6] described, so that they may be different subtypes of disease. In grade III gliomas, it has been already shown that patients with mutated *IDH1/2* carry 1p/19q codeletion, *TP53* gene mutation, and O⁶-methylguanine methyltransferase gene (*MGMT*) promoter methylation, while those with wild-type *IDH1/2* carry gain of 7p (*EGFR*), loss of 10q (*PTEN*), and homozygous deletion of 9p (*CDKN2A*) [3, 7, 8].

Patients with grade III gliomas present relatively poor overall survival (OS), ranging from 30 to 70 months [9, 10], and *IDH1/2* mutations were frequently found, ranging from 62 to 78% [1]. Therefore, patients in this grade obtain the most benefit from detailed analysis of *IDH1/2* gene status compared to patients with grade II glioma, who carry 71–78% of *IDH1/2* mutations [1] but present favorable prognosis (OS: over 100 months) [11, 12], and patients with primary glioblastoma, who carry only 7% of *IDH1/2* mutations [1]. Here, we conducted a retrospective study targeting detailed analysis of patients with grade III glioma. They were divided into two groups, mutated and wild-type *IDH1/2*, and each subgroup was investigated regarding associated prognostic factors. The factors analyzed were age, sex, preoperative Karnofsky performance status (KPS), extent of resection, oligodendroglial component, *MGMT* promoter methylation, Ki-67 labeling index, *TP53*, 1p/19q codeletion, and alteration of 7p (*EGFR*), 9p (*CDKN2A*), and 10q (*PTEN*), all of which are known prognostic factors of high grade gliomas. This study contains the largest number of analyzed factors, and we hypothesized that this subgroup analysis would enable us to clarify prognostic markers in either mutated or wild-type *IDH1/2* groups of grade III gliomas.

Patients and methods

Patients and samples

All patients underwent radical surgery and postsurgical treatment at Tohoku University Hospital. After the second review by a neuropathologist (H.S.), 8 patients were excluded and in total 115 grade III glioma patients [anaplastic astrocytoma (AA) 56, anaplastic oligoastrocytoma (AOA) 28, anaplastic oligodendroglioma (AO) 31] were analyzed for this study. Clinical profiles of each patient were obtained from the medical records. The resection rate was evaluated by gadolinium-enhanced magnetic resonance imaging (MRI) within 3 days after operation. The postsurgical treatment was chemotherapy [nimustine hydrochloride (ACNU) or temozolomide (TMZ)] and radiation therapy. After the initial treatment, patient observation

was performed by enhanced MRI every 2 months for the first 2 years, and every 3–6 months thereafter. Recurrence was judged by progression or new enhanced lesion on MRI. At recurrence, salvage surgery, second line chemotherapy or radiation therapy was performed. Tumor specimens were immediately frozen in liquid nitrogen and stored at -80°C until the extraction of DNA. This retrospective study was conducted with the approval of the ethics committee of Tohoku University School of Medicine, and written informed consent was obtained from all patients.

Outcome

The starting point of progression-free survival (PFS) and overall survival (OS) was the day of surgery. Tumor progression at last follow-up examination or death, and the last follow-up examination or death, were the end points for PFS and OS, respectively.

Ki-67 labeling index

The resected specimens were examined by immunohistochemical staining for Ki-67 antigen (Ki-67 antibody; DAKO, Tokyo, Japan). Each slide was individually reviewed and scored by one neuropathologist (M.W.). The Ki-67 labeling index was established by determining the percentage of positive nuclei among 1000 tumor cells. If the tumor had several areas of staining with different percentages, the highest percentage was applied for the analysis.

PCR amplification and direct sequencing

Genomic DNA extraction was carried out using the QIA-amp DNA Mini Kit (Qiagen Science, Germantown, Maryland, USA) following the manufacturer's protocol. Exon 4 of the *IDH1* and *IDH2* genes, and exons 4–9 of the *TP53* gene were amplified by PCR as previously described [13, 14]. The primer set of *IDH1/2* was as follows: *IDH1* sense 5'-CGG TCT TCA GAG AAG CCA TT-3' and antisense 5'-GCA AAA TCA CAT TAT TGC CAA C-3'; *IDH2* sense 5'-CAA GCT GAA GAA GAT GTG GAA-3' and antisense 5'-CAG AGA CAA GAG GAT GGC TA-3'. All PCR products were purified using a high pure PCR product purification kit (Roche, Basel, Switzerland) before the sequencing reaction. Sequencing analysis was as previously reported [13].

MGMT promoter methylation analysis

Methylation-specific polymerase chain reaction (MSP) was carried out for the *MGMT* gene promoter methylation analysis as previously described [15, 16]. The primer

sequences for methylation were 5'-TTT CGA CGT TCG TAG GTT TTC GC-3' (sense) and 5'-GCA CTC TTC CGA AAA CGA AAC G-3' (antisense), and for unmethylation were 5'-TTT GTG TTT TGA TGT TTG TAG GTT TTT GT-3' (sense) and 5'-AAC TCC ACA CTC TTC CAA AAA CAA AAC A-3' (antisense). MSP products were separated on 4% agarose gel containing ethidium bromide and the bands were visualized with ultraviolet light. Each sample was examined three times.

Detection of DNA copy number changes

Detection of chromosome copy number changes of 1p, 7p (*EGFR*), 9p (*CDKN2A*), 10q (*PTEN*), and 19q was carried out by multiple ligation-dependent probe amplification (MLPA) using a SALSA kit: P088 and kit: P105 probe mix (MRC Holland, Amsterdam, the Netherlands) as previously reported [17–19]. The former probe kit contains 15 1p probes, 8 19q probes, and 15 reference probes, and the latter kit contains 11 *PTEN* probes, 5 *CDKN2A* probes, 8 *TP53* probes, 2 *ERBB2* probes, 11 *EGFR* probes, and 8 reference probes. This technique consists of denaturing, hybridization (over 16 h at 60°C), ligation, amplification, and fragment analysis. Fragment analysis with the CEQ 2000 XL DNA analysis system (Beckman Coulter, Inc., Brea, CA, USA) provides subjective and objective data. Three control samples from healthy humans were kindly provided by FALCO Biosystems (Kyoto, Japan).

For data analysis, each probe amplification product is divided by the average of reference probes to compensate the difference in PCR efficiency, using Microsoft Excel. The data obtained are divided by the corresponding average probe fraction of the normal sample. Above 1.2 or below 0.8 is considered as gain or loss, respectively. In particular, below 0.4 is considered as homozygous deletion [17–20]. A truncated variant of the *EGFR* gene was analyzed as previously reported [18].

Statistical analysis

The relationship between *IDH1/2* gene status and other factors was evaluated by the Mann–Whitney and Fisher's exact tests. Probabilities of PFS and OS were calculated according to the Kaplan–Meier method and compared with the log-rank test. Genetic alterations, together with histology (histological types and Ki-67 labeling index), demographic (age and sex), clinical (preoperative KPS and surgery), and therapeutic variables achieving $P < 0.10$ in univariate analysis, were subsequently introduced in a backward stepwise proportional hazard analysis (Cox model) as independent predictors of survival. Variables which did not achieve $P < 0.10$ were listed as “not applicable (N/A)” (Tables 3, 5). All statistical methods

adopted a significance level of $P < 0.050$, using a statistical software package (SPSS, Inc., Chicago, IL, USA).

Results

Patients

Gross total resection was achieved in 48 patients, subtotal in 50 patients, partial in 17 patients and no biopsy. The postsurgical treatment was chemotherapy and radiotherapy in 95 patients (ACNU in 85, TMZ in 10), radiotherapy only in 12, and ACNU chemotherapy only in 5. The other three patients were treated with surgical resection only. There was no significant survival difference in patients treated with ANCU and TMZ (data not shown).

IDH1/2 gene mutation and other prognostic factors

The *IDH1/2* gene mutation was detected in 76 (66%) (70 *IDH1* and 6 *IDH2* gene mutations) of the 115 grade III glioma patients. Frequencies of mutations were 63% (AA), 75% (AOA), and 65% (AO). Among these, 39 patients with *IDH1/2* gene mutation were previously reported [13]. Age, sex, and resection rate were not correlated with *IDH1/2* status (Table 1). The Ki-67 labeling index was analyzed in 87 of the 115 patients. This proliferation index was lower in tumors with mutated *IDH1/2* than in those with wild-type *IDH1/2*, but without significant difference (17.0% and 23.7%, respectively, $P = 0.16$, Table 1).

The *TP53* gene mutation was detected in 51 (44%) of the 115 patients; 36/56 (64%) in AA, 10/28 (36%) in AOA, and 5/31 (16%) in AO. The frequency was 47% in patients with mutated *IDH1/2* and 38% in patients with wild-type *IDH1/2* ($P = 0.43$, Table 1). The *TP53* gene statuses of 22 cases were previously reported [21]. The 1p/19q codeletion was detected in 34 (30%) of 115 patients; 1/56 (1.8%) in AA, 14/28 (50%) in AOA, and 19/31 (61%) in AO. The frequency was 39% in patients with mutated *IDH1/2* and 10% in those with wild-type *IDH1/2* ($P = 0.0011$, Table 1). The 1p/19q statuses of 31 cases by fluorescent in-situ hybridization analysis were previously reported [22] and the same results were obtained by MLPA. One AO patient with *IDH1/2* gene mutation had both *TP53* gene mutation and 1p/19q codeletion, whereas 11 patients with *IDH1/2* gene mutation had neither *TP53* gene mutation nor 1p/19q codeletion (data not shown). *MGMT* promoter methylation was found in 87 (76%) of the 115 patients; 41/56 (73%) in AA, 22/28 (79%) in AOA, and 24/31 (77%) in AO. *MGMT* promoter methylation was significantly associated with *IDH1/2* gene mutation ($P < 0.0001$, Table 1). 7p gain (*EGFR* amplification) and 9p homozygous deletion were detected in 21 (18%) and 9 (7.8%) of

Table 1 Correlation between clinical/genetic factors and *IDH1/2* mutation

Clinical/genetic factor	Total <i>n</i> = 115	Mutated <i>IDH1/2</i> <i>n</i> = 76 (66%)	Wild-type <i>IDH1/2</i> <i>n</i> = 39 (34%)	<i>P</i>
Median age at diagnosis, year (range)	46 (10–77)	41 (10–70)	52 (14–77)	0.14 ^a
Sex, female, <i>n</i> (%)	45 (39)	34 (45)	11 (28)	0.11 ^b
Preoperative KPS \geq 80%, <i>n</i> (%)	91 (79)	62 (82)	29 (74)	0.47 ^b
Gross total resection, <i>n</i> (%)	48 (42)	30 (39)	18 (46)	0.55 ^b
Ki-67 labeling index, %	19.4	17.0	23.7	0.16 ^a
<i>TP53</i> mutation, <i>n</i> (%)	51 (44)	36 (47)	15 (38)	0.43 ^b
1p/19q codeletion, <i>n</i> (%)	34 (30)	30 (39)	4 (10)	0.0011^b
7p (<i>EGFR</i>) amplification, <i>n</i> (%)	21 (18)	8 (11)	13 (33)	0.0045^b
9p (<i>CDKN2A</i>) homozygous deletion, <i>n</i> (%)	9 (7.8)	3 (3.9)	6 (15)	0.060 ^b
10q (<i>PTEN</i>) loss, <i>n</i> (%)	10 (8.7)	6 (7.9)	4 (10)	0.73 ^b
<i>MGMT</i> gene promoter methylation, <i>n</i> (%)	87 (76)	69 (91)	18 (46)	<0.0001^b

KPS Karnofsky performance status

Bold indicates statistical significance at $P < 0.050$

^a Mann–Whitney test

^b Fisher's exact test

115 cases, respectively. The frequencies of 7p gain were 15/56 (27%) in AA, 6/28 (21%) in AOA, and none in AO, and of 9p homozygous deletion were 6/56 (11%) in AA, none in AOA, and 3/31 (9.7%) in AO. One AOA patient with *IDH1/2* gene mutation harbored a truncated variant of the *EGFR* gene. 7p gain (*EGFR* amplification) was significantly correlated with wild-type *IDH1/2* ($P = 0.0045$, Table 1). 9p homozygous deletion tended to be associated with wild-type *IDH1/2*, but did not reach statistical significance ($P = 0.060$, Table 1). 10q loss was found in 10 (8.7%) of the 115 patients; 7/56 (13%) in AA, 3/28 (11%) in AOA, and none in AO, without correlation with *IDH1/2* gene mutation ($P = 0.73$, Table 1).

Clinical/genetic prognostic factors were also analyzed by *MGMT* promoter methylation status and chromosome 1p19q status (see Electronic Supplementary Material Tables 1 and 2). *MGMT* promoter methylation was significantly associated with 1p19q codeletion ($P = 0.0003$, Supplementary Tables 1 and 2).

Progression-free survival and overall survival

The median follow-up was 43 months, ranging from 1 to 245 months. 62 patients (54%) remained free from progression. Median PFS was 62 months and 5-year PFS rate was 50.1%. Univariate analysis showed that the factors associated with longer PFS were young age (<50 years) ($P = 0.019$), low Ki-67 labeling index ($<15\%$) ($P = 0.014$), *IDH1/2* gene mutation ($P < 0.0001$), 1p/19q codeletion ($P = 0.0076$), no gain of 7p ($P = 0.0006$), and *MGMT* gene promoter methylation ($P < 0.0001$) (Table 2).

Multivariate Cox regression analysis revealed that *IDH1/2* mutation [hazard ratio (HR) 0.11, 95% confidence interval (CI) 0.056–0.23, $P < 0.0001$], low Ki-67 labeling index ($<15\%$) (HR 0.29, 95% CI 0.14–0.57, $P = 0.0004$), and no gain of 7p (HR 0.31, 95% CI 0.14–0.68, $P = 0.0034$) were independent favorable prognostic factors for PFS (Table 3), while 1p/19q codeletion and *MGMT* gene promoter methylation were not independent prognostic factors for PFS ($P = 0.59$ and 0.21, respectively, data not shown). 77 patients (67%) remained alive. Median OS was not reached and 5-year OS rate was 65.0%. Univariate analysis revealed that female sex ($P = 0.026$), gross total resection ($P = 0.049$), *IDH1/2* gene mutation ($P < 0.0001$) (Fig. 1), 1p/19q codeletion ($P = 0.0092$), no gain of 7p ($P < 0.0001$), and *MGMT* gene promoter methylation ($P = 0.033$) were correlated with longer OS (Table 2). Multivariate Cox regression analysis showed *IDH1/2* mutation (HR 0.16, 95% CI 0.073–0.37, $P < 0.0001$), low Ki-67 labeling index ($<15\%$) (HR 0.24, 95% CI 0.10–0.58, $P = 0.0014$), and no gain of 7p (HR 0.26, 95% CI 0.11–0.60, $P = 0.0018$) were independent favorable prognostic factors for OS (Table 3), but that 1p/19q codeletion and *MGMT* gene promoter methylation were not ($P = 0.63$ and 0.35, respectively, data not shown).

We observed relatively longer PFS and OS (Table 2) compared to previous reports [17]. This is largely due to the high resection rate of our series. In grade III gliomas, it is reported that radical resection is an independent better prognostic factor [17]. We excluded patients who underwent biopsy, and the extent of resection more than subtotal reached about 85%.

Table 2 Clinical and genetic parameters affecting PFS and OS in grade III gliomas

Characteristic	PFS				OS			
	No. (%)	5-year survival (%)	Median months	<i>P</i> (log-rank)	No. (%)	5-year survival (%)	Median months	<i>P</i> (log-rank)
	115	50.1	62		115	65.0	NR	
Histology								
AA	56 (49)	44.3	60	0.19 (vs AO, AOA)	56 (49)	58.3	88	0.092 (vs AO, AOA)
AO	31 (27)	54.1	62		31 (27)	79.5	NR	
AOA	28 (24)	57.5	NR		28 (24)	65.5	NR	
Age at diagnosis								
<50 years old	67 (58)	59.0	NR	0.019	67 (58)	67.0	NR	0.35
≥50 years old	48 (42)	36.1	47		48 (42)	61.8	81	
Sex								
Female	45 (39)	58.4	NR	0.077	45 (39)	79.3	NR	0.026
Male	70 (61)	44.6	47		70 (61)	54.8	88	
Karnofsky performance status								
≥80	91 (79)	51.3	68	0.30	91 (79)	64.4	NR	0.22
<80	24 (21)	46.4	50		24 (21)	65.7	62	
Gross total resection								
Yes	48 (42)	59.3	NR	0.094	48 (42)	71.8	NR	0.049
No	67 (58)	42.5	51		67 (58)	60.0	81	
Ki-67 labeling index								
<15%	39 (45)	59.7	95	0.014	39 (45)	71.5	NR	0.090
≥15%	48 (55)	32.9	38		48 (55)	56.0	NR	
IDH1/2								
Mutated	76 (66)	70.5	NR	<0.0001	76 (66)	82.3	NR	<0.0001
Wild-type	39 (34)	10.1	11		39 (34)	26.3	26	
TP53								
Wild-type	64 (56)	53.2	95	0.46	64 (56)	70.0	NR	0.13
Mutated	51 (44)	46.0	54		51 (44)	58.5	88	
1p/19q codeletion								
Yes	34 (30)	69.4	NR	0.0076	34 (30)	85.0	NR	0.0092
No	81 (70)	42.1	35		81 (70)	56.2	88	
Gain of 7p								
No	94 (82)	57.4	95	0.0006	94 (82)	74.9	NR	<0.0001
Yes	21 (18)	15.1	18		21 (18)	15.6	27	
Homozygous deletion of 9p								
No	106 (92)	52.3	68	0.17	106 (92)	66.1	NR	0.080
Yes	9 (8)	27.8	22		9 (8)	53.3	64	
Loss of 10q								
No	105 (91)	51.9	68	0.22	105 (91)	66.2	NR	0.31
Yes	10 (9)	32.0	17		10 (9)	56.0	NR	
MGMT								
Methylated	87 (76)	57.2	NR	<0.0001	87 (76)	69.6	NR	0.033
Unmethylated	28 (24)	24.8	15		28 (24)	44.5	43	

AA anaplastic astrocytoma, AO anaplastic oligodendroglioma, AOA anaplastic oligoastrocytoma, NR not reached, OS overall survival, PFS progression-free survival

Bold indicates statistical significance at $P < 0.050$

Table 3 Multivariate analysis of independent factors associated with survival in grade III glioma patients

Variable	PFS			OS		
	Hazard ratio	95% CI	P	Hazard ratio	95% CI	P
<i>IDH1/2</i>						
Mutated versus wild-type	0.11	0.056–0.23	<0.0001	0.16	0.073–0.37	<0.0001
Histological diagnosis						
AO, AOA, vs AA	N/A	N/A	N/A	0.47	0.20–1.12	0.088
Ki-67 labeling index						
<15 vs ≥15%	0.29	0.14–0.57	0.0004	0.24	0.10–0.58	0.0014
7p (<i>EGFR</i>)						
Gained (–) vs (+)	0.31	0.14–0.68	0.0034	0.26	0.11–0.60	0.0018

N/A not applicable, OS overall survival, PFS progression-free survival, CI confidence interval

Bold indicates statistical significance at $P < 0.050$

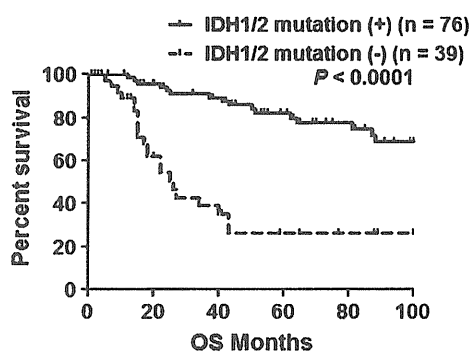


Fig. 1 Overall survival of patients with grade III gliomas by isocitrate dehydrogenase 1 and 2 (*IDH1/2*) gene status. Patients with mutated *IDH1/2* present prolonged survival compared with patients with wild-type *IDH1/2*

Subclassification of Grade III Gliomas Based on *IDH1/2* Status

Patients were divided into two groups by *IDH1/2* gene status. In the mutated *IDH1/2* group, median PFS and OS were not reached, and 5-year survival rate was 70.5 and 82.3%, respectively. Gross total resection was associated with better PFS ($P = 0.040$; Table 4), and gross total resection, wild-type *TP53* gene, and no gain of 7p were associated with longer OS ($P = 0.018$, 0.046, and 0.017, respectively) (Table 4; Fig. 2a–c). Multivariate Cox regression analysis in the mutated *IDH1/2* group revealed that 7p gain (HR 5.3, 95% CI 1.35–20.5, $P = 0.017$) and high Ki-67 labeling index ($\geq 15\%$) (HR 3.2, 95% CI 1.11–9.05, $P = 0.030$) were independent prognostic factors of poor PFS (Table 5), and incomplete resection (HR 5.7, 95% CI 1.53–20.8, $P = 0.0092$), 7p gain (HR 5.5, 95% CI 1.39–21.5, $P = 0.015$), and *TP53* gene mutation (HR 3.4, 95% CI 1.15–9.84, $P = 0.026$) were independent prognostic factors of poor OS (Table 5).

In the wild-type *IDH1/2* group, median PFS and OS were 13 and 26 months, respectively, and 5-year survival

rate was 13.5 and 26.3%, respectively. Low Ki-67 labeling index ($<15\%$) was associated with longer PFS ($P = 0.049$; Table 6) by univariate analysis. Multivariate analysis revealed high Ki-67 labeling index ($\geq 15\%$) (HR 2.6, 95% CI 1.05–6.16, $P = 0.038$) and 10q loss (HR 3.7, 95% CI 1.01–13.6, $P = 0.047$) were independent prognostic factors of poor PFS (data not shown). For OS in the wild-type *IDH1/2* group, no factor showed statistical significance, but oligodendroglial component and 1p19q codeletion tended to be associated with longer OS by univariate analysis ($P = 0.094$, $P = 0.099$, respectively) (Table 6).

Additionally, grade III patients were divided into two groups by either *MGMT* gene promoter status or chromosome 1p/19q status to see the prognostic effects of *IDH1/2* mutation on each group. Mutated *IDH1/2* subgroups presented better survival in all groups except for 1p19q codeletion & *IDH1/2* wild-type, which did not reach statistical significance due to the small number of patients ($n = 4$) ($P = 0.74$, Supplementary Fig. 1).

Discussion

We mainly focused on the subgroup analysis of grade III glioma based on *IDH1/2* gene status. *TP53* mutation has obtained the status of a hallmark of gliomas [23], but remains a controversial prognostic factor. In one report, *TP53* gene mutation was proposed as a factor in prolonged survival in patients with AA and glioblastoma [24], while in other, *TP53* gene mutation in AA, AOA, and AO was found to indicate shorter survival [25]. Interestingly, we observed that in patients with grade III glioma in the mutated *IDH1/2* group the absence of *TP53* gene mutation was correlated with longer survival (PFS $P = 0.053$, OS $P = 0.046$) (Table 4; Fig. 2b), and was an independent favorable prognostic factor of OS by