Int. J. Radiation Oncology Biol. Phys., Vol. 56, No. 2, pp. 436–447, 2003

0360-3016/03/$–see front matter

doi:10.1016/S0630-3016(02)04612-6

EXPERIENCE OF ULTRASOUND-BASED DAILY PROSTATE LOCALIZATION

ANURAG CHANDRA, M.D.,* LEI DONG, PH.D.,† EUGENE HUANG, M.D.,* DEBORAH A. KUBAN, M.D.,* LAURA O'NEILL, M.S.,† ISAAC ROSEN, PH.D.,† AND ALAN POLLACK, M.D., PH.D.*‡

Departments of *Radiation Oncology and †Radiation Physics, The University of Texas M. D. Anderson Cancer Center, Houston, TX; ‡Department of Radiation Oncology, Fox Chase Cancer Center, Philadelphia, PA

Purpose: The NOMOS (Sewickley, PA) B-mode Acquisition and Targeting System (BAT) ultrasound system provides a rapid means of correcting for interfraction prostate positional variation. In this investigation, we report our experience on the clinical issues relevant to the daily use of the BAT system and the analysis of combined setup error and organ motion for 3509 BAT alignment procedures in 147 consecutive patients treated with IMRT for prostate cancer.

Methods and Materials: After setup to external skin marks, therapists performed the BAT ultrasound alignment procedure before each IMRT treatment. In this study, a single physician (A.C.) reviewed all BAT images and classified image quality and accuracy of image alignment by the therapist. On a scale of 1–3, near-perfect image quality or alignment was given a 1, fair image quality or misalignment ≤5 mm (likely within the PTV) was given a 2, and unacceptable image quality or misalignment >5 mm (potential to violate the PTV) was given a value of 3. The distribution of shifts made was analyzed in each dimension and for all patients. The time required to perform the BAT alignment was also assessed in 17 patients.

Results: Among the 3509 attempted BAT procedures, the image quality was judged to be poor or unacceptable in 5.1% (181). Of the remaining 3328 BAT images, with quality scores of 1–2, alignments were unacceptable (>5 mm misalignment as judged by the reviewing physician) in 3% (100). The mean shift in each direction, averaged over all patients, was 0.5–0.7 mm. Interfraction standard deviation (1 SD) of prostate position based on combined setup error and internal organ motion is 4.9 mm, 4.4 mm, and 2.8 mm in the anteroposterior (AP), superior-inferior (SI), and lateral (RL) dimensions, respectively. The distribution of the shifts was a near-random Gaussian-type in all three major axes, with greater variations in AP and SI directions. The percent of BAT procedures in which the shift was >5 mm was 28.6% in AP, 23% in SI, and 9% in RL directions. The average BAT procedure took extra 5 min out of a 20-min time slot in a typical eight-field IMRT treatment.

Conclusion: The quality of the daily ultrasound images was deemed acceptable in 95%. Major alignment errors by therapists were only 3%. The BAT system is clinically effective and feasible in a matter of 5 min. Although the accuracy of the BAT was not addressed in this investigation, we found a significant percentage of large shifts being made from the initial alignment position.

Prostate, Localization, Ultrasound, Radiation, Alignment.

INTRODUCTION

Prostate cancer treatment with radiotherapy has dramatically changed over the last 10 years with the implementation of conformal, and, more recently, intensity-modulated radiation therapy (IMRT) techniques. Dose escalation is commonly being performed with the advent of conformal radiotherapy (1–7). To safely achieve prostate doses above 75 Gy, there is a greater demand to use tighter margins, especially posteriorly around the rectum to reduce complications (8). A large margin of 1.5 cm in the posterior direction results in significant rectal bleeding when delivering prostate doses above 70 Gy (8).

Studies have shown that standard deviation (SD) for setup variation is on the order of 2.0–3.0 mm (9–13), and that internal changes in prostate position generally vary on the order of 2.0 to 3.0 mm in the anteroposterior (AP) and superior-inferior (SI) dimensions and 1.0 mm in the lateral (RL) dimension (9, 11, 13–17). Combined setup and organ motion studies have shown general variability on the order of 4.0 to 6.0 mm in the SI and AP dimension and 3.0 to 5.0 mm in the RL dimension (11, 13, 18). Currently, margins used for the majority of patients are based on such population averages. Calculations by various authors have shown a planning target volume (PTV) margin of approximately 7

Reprint requests to: Lei Dong, Ph.D., Department of Radiation Physics, Box 94, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030. Tel: (713) 745-4502; Fax: (713) 745-4505; E-mail: ldong@mail.mdanderson.org

Acknowledgments—The authors would like to thank Mark F. Munsell of the Department of Biostatistics for helpful data analysis and Mohammad Salehpour, Ph.D., Department of Radiation Physics, for assisting in BAT data collection.

Received Jan 2, 2002, and in revised form Dec 12, 2002. Accepted for publication Dec 13, 2002.

mm in the RL dimension and 9–12 mm in the SI and AP dimensions is needed to ensure that the prostate, or clinical target volume (CTV), is included in the PTV 95% of the time (9, 10, 13, 14). Because such a margin can lead to untoward toxicity, many have resorted to reducing the PTV for part of the treatment (8). Reducing margins may compromise treatment efficacy unless setup variations are reduced or daily adjustments are made for interfraction changes in internal anatomy. Unless prostate position is monitored, there is no guaranty that the CTV will be covered adequately (19).

A number of techniques have been developed for measuring setup variations and internal organ motions for individual patients during a course of treatment and treatment margins are reduced based on these techniques. The first method developed for this purpose measures the patient's position using port films on a daily basis to establish the setup variation unique to that patient (9, 11–13). Subsequently, margins are reduced based on those individual-specific data. This method requires many port images and some data analysis, and does not address internal organ position changes. Another method requires an invasive procedure to implant fiducials in the prostate (11, 12). They can then be imaged on a daily basis to locate the position of the prostate before treatment and adjust the patient's location. Rectal balloons (13) and radio-opaque urethral catheters (14) can be used for both immobilization and imaging. The most widely used method for assessing prostate position on a daily basis and correcting for variations is the B-mode Acquisition and Targeting System (BAT) ultrasound system (18, 20–22) from NOMOS Corporation (Sewickley, PA).

The BAT ultrasound system provides direct visualization and targeting of the prostate on a daily basis before treatment. Prior studies by Lattanzi et al. (18) have shown that three-dimensional shifts described by the BAT system to correct for interfraction movement are correlated with those measured by computed tomography (CT) localization of the prostate. The study suggests that BAT imaging is comparable to CT imaging for targeting purposes. Although the range of differences between the same patient setup measured by the BAT and the CT is up to 7 mm, the averaged differences are small in 69 measurements (18). Definitive assessment of the accuracy of the BAT itself is not addressed in our study.

In this study, we investigated operational issues relevant to the daily use of the BAT in a busy clinic and analyzed the pattern of positioning variations for combined setup error and organ motion from daily BAT measurements. The clinical benefit of the BAT was dependent on the skill of the therapists in obtaining meaningful images and aligning those images with the planning contours obtained from the simulation CT scan. We studied the quality of the BAT images, the accuracy of alignment by the therapists, and the impact on clinic schedule and linac utilization. More than 3500 BAT alignment images were evaluated for this purpose. Also, 3228 alignment shifts were analyzed, making this to be one of the largest studies of combined setup error and organ motion using the BAT ultrasound system.

METHODS AND MATERIALS

Patient and treatment characteristics

Between December 1999 and December 2000, 147 consecutive patients were treated with IMRT using the BAT ultrasound system at the University of Texas M. D. Anderson Cancer Center. IMRT was given as a boost (median 17 fractions) in 79 patients or as full treatment (42–44 fractions) in 68 patients. There were 4158 IMRT fractions administered and 3509 BAT alignment procedures performed. On the first day of treatment, BAT alignment was intentionally not performed so that the patient was given enough time for treatment instructions and taking orthogonal portal films. Occasionally, a BAT alignment was not done because of technical problems with the linear accelerator or unexpected difficulties in the patient's treatment schedule, necessitating treatment on an alternative therapy machine not equipped with the BAT. Technical problems with the BAT itself were rare.

CT simulation and treatment planning

All patients underwent a planning simulation CT scan with a full bladder and empty rectum. This was accomplished by having the patient drink 16 to 20 oz of fluid 30–60 min before the CT scan. An enema was used before the simulation to empty the rectum. To minimize setup variability, the patient was placed in a VAC-LOK (Med-Tec, Inc., Orange City, IA) cradle from the mid-thigh to the feet. The CT scan was then obtained in the supine treatment position. Isocenter was placed in the prostate and marked on the external skin with tattoos.

Using the initial planning CT scan, an IMRT treatment plan was designed on an inverse planning system (Corvus, NOMOS Corp.). Treatment was delivered using either serial tomotherapy via a binary multileaf collimator (MIMiC, NOMOS Corp.) or a segmental multileaf collimator (SMLC) technique (Varian Associate, Palo Alto, CA) with fixed gantry angles. Patients had either a 17-fraction boost treatment or full treatment with 42 or 44 fractions at 1.8 Gy per fraction. The planning margin was 10 mm in anterior and lateral directions, 5–8 mm in superior and inferior directions, and 4–6 mm in the posterior direction. The current margins used were adopted from a previous, repeated CT study by Antolak et al. (14).

BAT ultrasound localization of the prostate

The BAT system consists of a B-mode transabdominal ultrasound probe attached to a precision tracking arm. This arm, along with computer software, allows the ultrasound image to be calibrated to the coordinates of the treatment machine. The ultrasound probe displays real-time images on the computer screen. The ultrasound probe is oriented to a reference point through an external docking system for the particular treatment room used. Thus the probe can be maneuvered in all directions and remain oriented to the isocenter because it recognizes its position in three-dimensional (3D) space relative to the docking system.

With the patient in treatment position, BAT ultrasound localization of the prostate is performed immediately before IMRT treatment. A trained radiation therapist captures both sagittal and transverse images. Patient contours (prostate, seminal vesicles, bladder, and rectum) from the planning CT scan are superimposed on the BAT images using the established isocenter from the docking system to relate the two coordinate systems. The therapist then corrects any misalignment between the CT contours and the BAT images. This is done by maneuvering the transverse and sagittal CT contours with a touch screen menu to precisely superimpose the ultrasound images with the CT contours originally transferred from the patient's treatment plan. When the images are aligned on the monitor, the computer reveals the couch shifts in three dimensions needed to bring the prostate into alignment with the original planning CT position.

Daily setup technique

Patients were treated in supine position with the vac-lok cradle. Initially, the isocenter was aligned to the simulation tattoos. On the first treatment day, orthogonal portal films were taken. Then, at least weekly portal films were taken to ensure that the skin marks were a good representation of internal bony landmarks. If the portal film indicated a relatively large shift (usually greater than 5 mm), a shift will be made and the patient's skin marks will be redrawn based on the portal film. Then the BAT alignment is subsequently made from the new skin marks. On the other 4 days of the week, the patient was initially aligned to skin marks, and then shifts were made as per the BAT procedure. The patient was always treated based on the shift detected by the BAT procedure. On the day when a BAT alignment was not available, patients were aligned using skin marks or based on portal films if portal filming was scheduled for that day.

In our clinical use, whenever a 3 mm or greater couch shift was made in any one direction, a repeat confirmatory BAT alignment was performed after the initial couch shift. If greater than 1-cm shift was required in any direction, the attending physician was notified and the whole setup procedure was repeated. Consistent shifts of more than 1 cm (5 patients in our study group) led to a second CT simulation for those patients, and thus a new treatment plan was used to do the BAT on these patients (data not shown).

A positive sign in right-left (RL) direction means the couch was shifted to the right. A positive sign in the AP direction means the couch was moved posteriorly (down), and a positive sign in SI direction means the couch was moved superiorly (moved toward the gantry).

BAT image review

In this study, a single physician (A.C.) reviewed all the BAT images. Image quality and the accuracy of image alignment done by the therapist were individually evaluated in each direction using a 1–3 scale. Images were scored as 1 for near-perfect quality, 2 for fair quality, and 3 for unacceptable. Therapist alignment was analyzed for images with good or fair quality (scores of 1 or 2) only. Images with unacceptable quality were excluded. Accurate alignment by the therapist (within 2 mm) was scored a 1; a minor misalignment of 2 mm but 5 mm was scored a 2; and a major misalignment of 5 mm was scored a 3. A major misalignment indicated a potential miss of the CTV. For major (5 mm) misalignments, the directions of misalignment were also recorded.

Couch shift analysis

Couch shifts were studied for all treatments where both the image quality and alignment were fair or better (scores of 1–2). These couch shifts were analyzed for their mean and variation in each direction and their changes over time. A corresponding 95% confidence interval (CI) was also computed. Systematic shifts in BAT alignment were assessed by aggregating all couch shifts to compute a mean shift in each direction. Group systematic shifts were estimated by combining individual systematic shifts over a population of patients. A correlation study was also performed among the shifts in three orthogonal directions.

We can calculate the spreads (in 1 SD) of the systematic shifts for the population of 147 patients used in this study. There were also random variations in couch shifts associated with a particular patient. We can calculate the individual random component from the series of alignment for the same patient and calculate the root-mean-square (RMS) of all SDs of 147 patients. The result of the systematic spread ( ) and random spread ( ) was to derive treatment margins for uncorrected shifts in prostate cancer treatment (23, 24). The margin recipe for covering 95% of the CTV for 90% of patients was proposed by Van Herk et al. (23) as 2.5 0.7 .

To determine whether there was a change in shift measurements over time, a linear model was fit to each measurement for each direction using the form Y X, where Y is an individual couch shift, X is the number of days from the date of the first measurement, and is the slope of the line. Weighted least squares were used to fit this model.

To assess whether the pattern of patient setup changed over time, the variance in couch shifts for an individual patient was computed and plotted at the date when the patient started his treatment. Then the variances for the group of patients were fitted to a linear model whose slope represented the change in variance over time. It was assumed that the variance in the population would not change over time.

Analysis of time factors

To assess the time required to complete a BAT alignment procedure, a time study was performed on 17 prostate cancer patients under treatment. The time was recorded when the patient entered the treatment room, when the first beam and last beam were completed, and when the room was ready for next patient. The average time for patient setup (including BAT), average treatment time per beam (including gantry rotation time and beam-on time), and patient dressing and room cleaning time were calculated. To derive the time required to perform the extra BAT alignment, the therapists used a timer to measure the time from the start to the end of the BAT procedure for 10 patients, from which the percentage of BAT time in overall treatment was derived.

RESULTS

Of the 3509 BAT procedures performed, image quality was acceptable (scores of 1 or 2) in 94.8% (3328 images). In 0.6% of procedures (22), no shift was made because the image was perceived to be too poor by the radiation therapist and therefore the patient was treated based on the bony anatomic isocenter without a BAT-based shift. In 4.5% of procedures (159), a shift was made, but, in retrospect, the image quality was judged to be unacceptable by the reviewing physician. Accuracy of alignments could not be assessed in the studies with unacceptable image quality.

Of the 3328 BAT studies with acceptable image quality, alignment by the therapist was unacceptable in 3.0% (100 images). These studies had major (5 mm) misalignments between the images and the contours from the treatment plan. The majority of these misalignments were in the superior/inferior direction (61%). Major misalignments in the anterior/posterior direction were seen in 55% of the studies. In the lateral direction, only 6% of studies had a major misalignment. Eleven cases (11%) had major misalignment in more than one dimension. An example of a good quality and an acceptable alignment is shown in Fig. 1a (scores of 1 for both categories). An example of good quality image (score of 1), but with the alignment judged unacceptable (score of 3), is shown in Fig. 1b. For the same patient, an example of poor image quality (score of 3) is shown in Fig. 1c.

Couch shifts were studied in the 3228 BAT images (92% of all BAT procedures) that had acceptable image quality (scores of 1–2) and acceptable alignment (scores of 1–2). Figure 2 shows the distribution of shifts in each direction and the composite shift along the vector of prostate motion (3D vector, Fig. 2 days). The distribution of shifts in each dimension is approximately symmetric and centered at zero, suggesting a random process. Figures 2a , 2b, and 2c clearly show that the variations in the AP and SI dimensions are larger than in the RL dimension. Figure 2 days shows that the majority of alignments have 3D shifts between 3.0 mm and 9.0 mm.

Table 1 summarizes the couch shifts. The mean shift in each direction was less than 1 mm. Although the shifts in the AP and SI directions are statistically different from zero, representing small systematic errors in overall procedure, the magnitudes of the systematic shifts (all less than 1 mm) are clinically inconsequential. The median magnitudes of the couch shifts in RL, SI, and AP directions were 1.1 mm, 2.5 mm, and 3.0 mm, respectively. The magnitude of the median 3D couch movement vector was 5.7 mm. The variance was greatest in the AP dimension (2.4 mm) compared with the SI dimension (1.9 mm), which was in turn greater than the RL (0.8 mm) dimension by the F test (p 0.001).

Figure 3 shows the distribution of average couch shift per patient. A nonzero averaged shift signifies a systematic offset from the marked skin tattoos to the actual treatment positions based on the BAT images. The nonzero systematic shift implies that the pretreatment CT scan on which the treatment plan is based is not typical of the patient's intreatment position. For the group of 147 patients in the study, the distribution appears to be random, approximately symmetric, and centered around zero. The greatest systematic shifts for individual patients were observed in the AP and SI directions. Table 2 summarizes these data. The mean values in Table 2 were calculated as the average of the mean shifts for individual patient, which are almost identical to the mean values calculated for individual shifts (Table 1). A small mean value implies that the procedure has few overall systematic errors.

The systematic shifts (in 1 SD) for the population of 147 patients used in this study are listed in Table 2 as . The random variations for the group of 147 (calculated using the RMS of individual random variations of 147 patients) are also listed in Table as . With the current study, treatment margins can be recommended for the conventional setup method using only external skin marks. Using the margin recipe proposed by Van Herk et al. (23), the recommended margins (for covering 95% of the CTV for 90% of patients) can be calculated as 5.3 mm, 9.2 mm, and 8.3 mm in the RL, AP, and SI directions, respectively. Alternatively, we can also calculate the margins in three directions using the overall alignment spreads in Table 1 (which includes both systematic and random components). Assuming a Gaussian distribution, 2-SD give approximately 95% coverage. The margins, thus calculated, are 5.6 mm, 9.8 mm, and 8.8 mm in RL, AP, and SI directions, which are similar to the margins estimated by the Van Herk's method.

To study any cross-correlation among the shifts in three directions, Fig. 4 shows the scatter plot of all alignments in AP/RL (Fig. 4a), SI/RL (Fig. 4b), and AP/SI (Fig. 4c) directions. The data in all three directions showed normal distribution (p 0.001) in the Kolmogorov-Smirnov normality test with Lilliefors correction (SPSS release 10, SPSS Inc., Chicago, IL). The normal Q-Q plots (not shown) also confirmed this normal distribution graphically. We found that there was no significant correlation between AP and RL directions (p 0.78); however, statistical correlation exists in SI/RL directions (p 0.001) and more strongly in AP/SI directions (p 0.001). The Pearson correlation coefficient (2-tailed) of 0.37 was weak but significant in AP/SI directions, with the minus sign indicating that a posterior shift is more likely associated with an inferior shift. This was also confirmed in the ranked correlation tests. The Spearman correlation (2-tailed) was

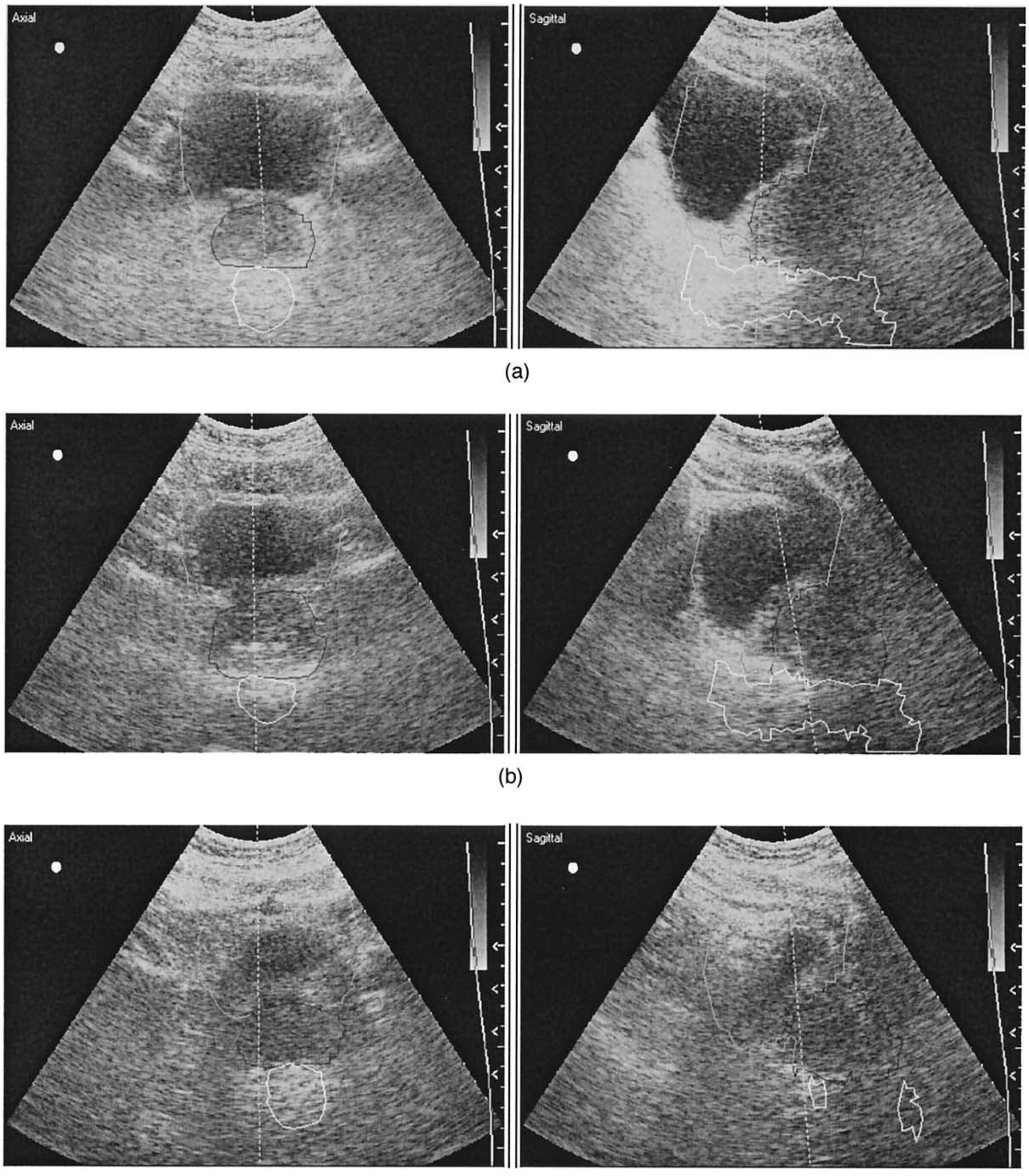

Fig. 1. (a) Example of a good image quality (score of 1) and acceptable alignment (score of 1); (b) example of a good image quality (score of 1) but unacceptable alignment (score of 3); and (c) example of a poor quality image (score of 3).

0.038 (p 0.033) in AP/RL, 0.048 (p 0.006) in SI/RL, and 0.396 (p 0.001) in AP/SI directions, respectively. A residual analysis and curve estimation confirmed a linear relationship in the AP/SI direction (analysis of variance F value 498, p 0.001). Figure 4d graphically shows that the couch shifts were most likely along the diagonal directions, as indicated by the arrow. In Fig. 4d, the prostate, bladder, and rectum were overlaid on a sagittal CT image along with typical IMRT isodose lines. Although not directly studied here, the AP/SI correlation suggests that the internal target (prostate) position was likely influenced by the combined effect of bladder and rectal volume variations. In this study, the target rotation can not be assessed because the BAT can correct translational shifts only.

The proportion of BAT studies in which there was a large shift (5 mm), with potential marginal miss of the CTV if uncorrected, is quantified in Table 3. Fully 28.6% of AP, 23.0% of SI, and 9.2% of RL alignments required a large

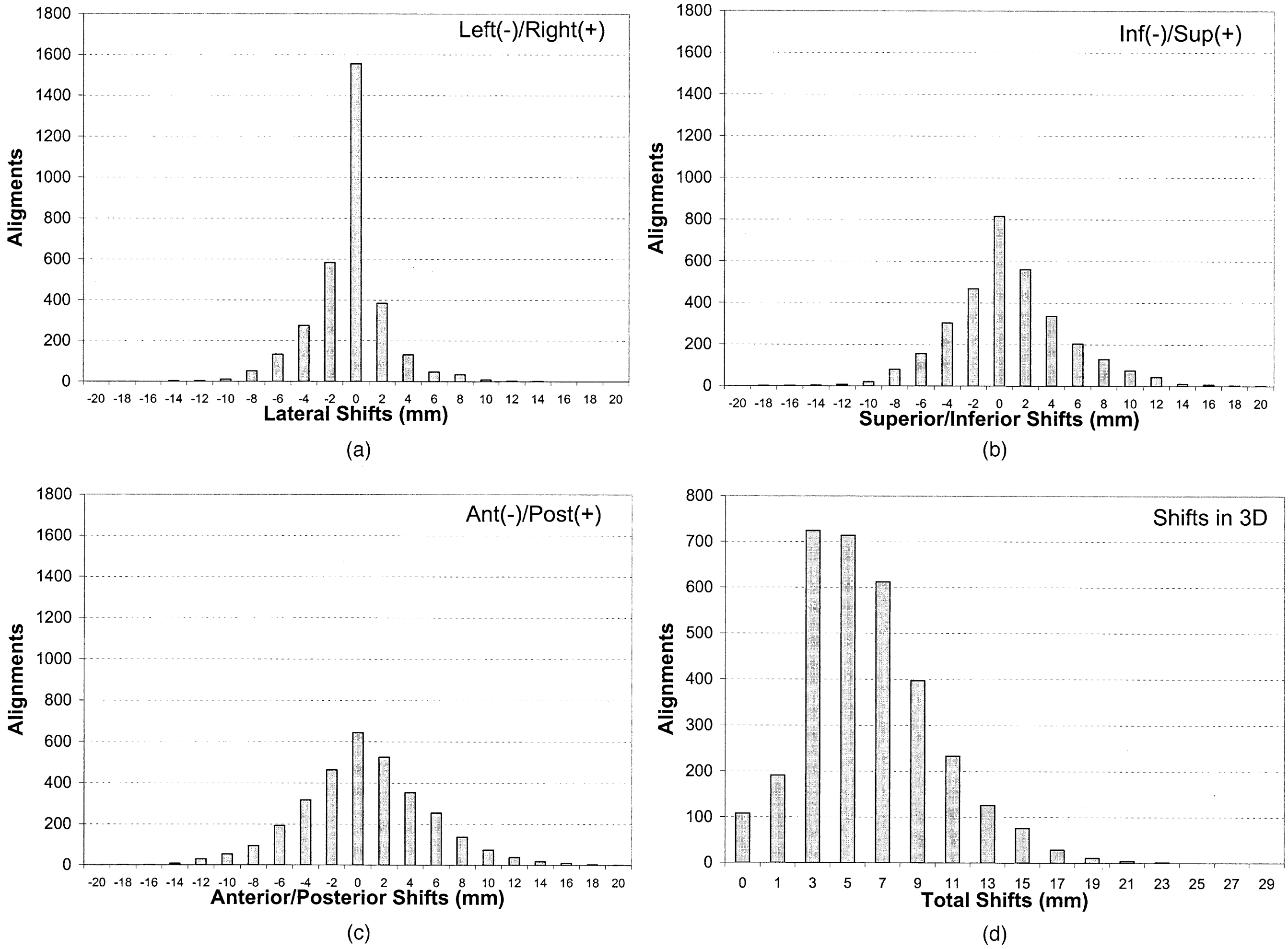

Fig. 2. Distribution of individual couch shifts using BAT relative to the initial skin-mark setup. A positive sign in the anteroposterior direction means the couch was moved posteriorly (down) and a positive sign in the superior-inferior direction means the couch was moved superiorly (toward gantry). The graphs show a random variability of shifts with the zenith around zero. (d) The amplitude of three-dimensional (3D) couch movement vector based on individual shifts in the three major axes; the median shifts (in 3D) are approximately 6 mm.

shift to adequately align the ultrasound image with the CT contours.

To investigate for any potential trends in using the BAT, couch shifts are plotted as a function of time in each dimension. Figure 5 shows the 3228 couch shifts in each direction plotted by their treatment date from December 1999 to December 2000. For each direction, individual shifts were fit to a linear model to determine whether there is a (linear) trend in these measurements over time. In the RL and SI directions, the slopes of the fit lines were not statistically different from zero (p 0.547, p 0.705, respectively), indicating that these mean shifts did not sig-

Table 1. Statistics for couch shifts for all alignments

Total alignments (3228)	RL (mm)	AP (mm)	SI (mm)	3D couch shifts (mm)*
Mean value (95% CI)	0.54 (0.64 to 0.45)	0.51 (0.34 to 0.68)	0.71 (0.56 to 0.86)	6.2
Spread (1 SD)	2.8	4.9	4.4	3.7
Median shifts	1.1	3.0	2.5	5.7
95% confidence range	6.7 to 5.7	9.5 to 10.6	7.8 to 10.6	0 to 13.2
Full range (extremes)	14 to 14	17 to 21	17 to 20	0 to 23

Abbreviations: RL lateral; AP anteroposterior; SI superior/inferior; 3D three-dimensional; CI confidence interval; SD standard deviation.

* Magnitude of three-dimensional couch movement vector.

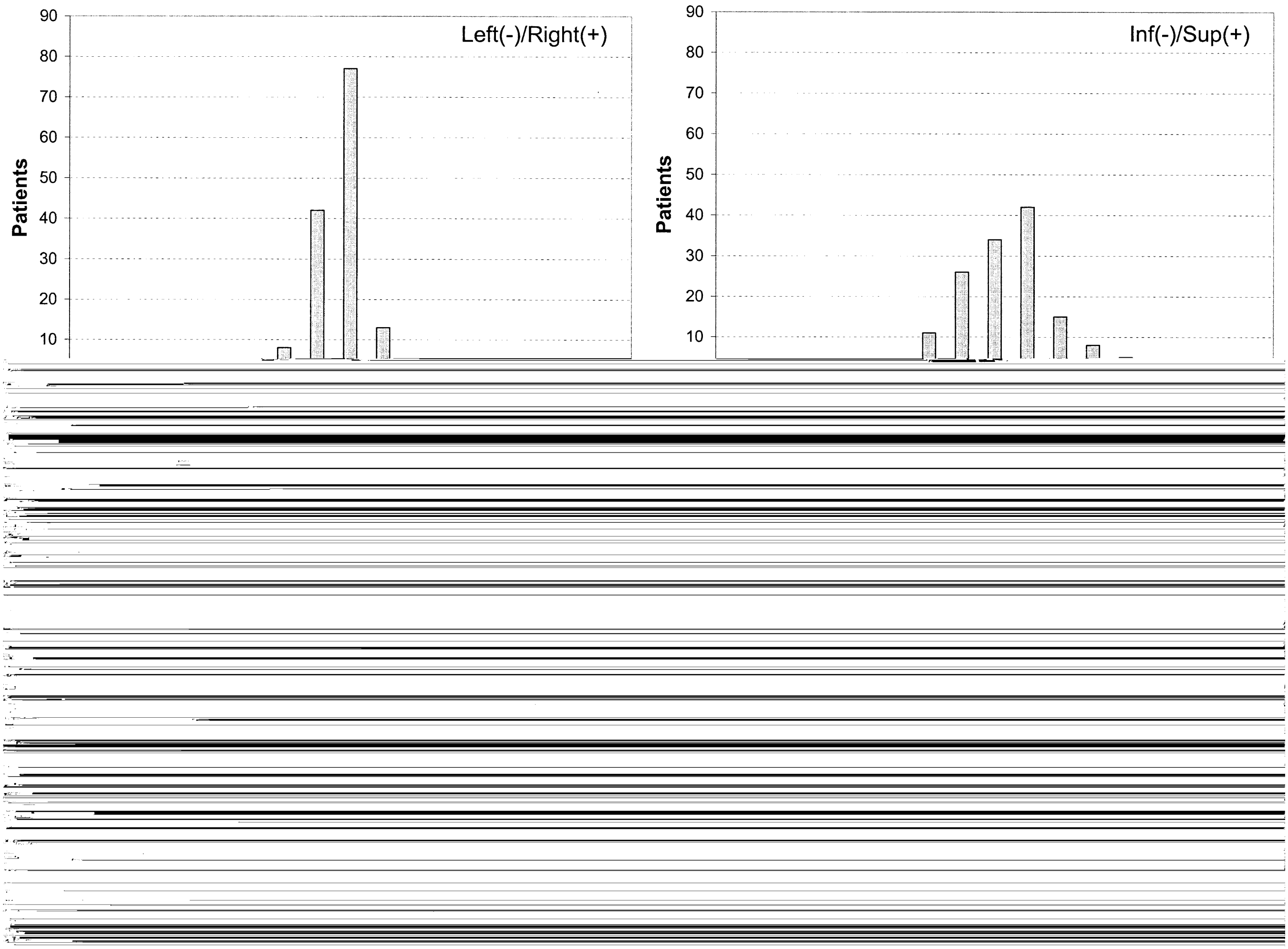

Fig. 3. Distribution of averaged couch shifts for each patient. The individual shifts for each patient were averaged for each dimension. The spread of shifts in the anteroposterior and superior-inferior dimensions is greater than the lateral dimension. (d) The amplitude of three-dimensional couch movement vector averaged for each patient.

nificantly change over time. In the AP direction, the slope of the fitted line was statistically different from zero (p 0.001), but the slope, 0.003 mm/day, is fairly small and likely not to be clinically significant for an individual patient in 42 fractions of treatment.

To analyze for any trends in the spread of shifts applied for an individual patient, Fig. 6 plotted the 2 SD, which is approximately the full range of shifts for a patient, during the time the patient started his treatment. Each of the 147 patients was graphed by his treatment start date, and a linear regression was done to obtain a linear fit for the 147 patients. Over time, the spread in shifts was not significantly changed in the RL dimension (p 0.1), but increased statistically in the SI and AP dimensions (p 0.001). This

Table 2. Statistics of couch shifts by patients

Total patients (147)	RL (mm)	AP (mm)	SI (mm)	3D couch shifts (mm)
Mean value*	0.49	0.59	0.92	Median 4.1
(1 SD) (systematic spread)	1.9	3.4	3.2	2.5
(RMS of random shifts)	2.3	3.3	2.7	2.9
95% confidence range	4.4 to 4.0	7.7 to 6.9	5.3 to 8.3	1.1 to 1.1
Full range (extremes)	6.0 to 9.4	8.4 to 10.6	6.4 to 10.7	0.7 to 12.6

Abbreviations: RL lateral; AP anteroposterior; SI superior/inferior; 3D-three-dimensional; SD standard deviation.
* The average of mean shifts for each individual patient for each dimension.