

● *Physics Original Contribution*

## HIGH-PRECISION PROSTATE CANCER IRRADIATION BY CLINICAL APPLICATION OF AN OFFLINE PATIENT SETUP VERIFICATION PROCEDURE, USING PORTAL IMAGING

ARJAN BEL, PH.D.,\* PIETER H. VOS, PH.D.,<sup>†</sup> PATRICK T. R. RODRIGUS, M.D.,<sup>‡</sup>  
CARIEN L. CREUTZBERG, M.D.,<sup>†</sup> ANDRIES G. VISSER, PH.D.,<sup>†</sup> JOEP C. STROOM, M.Sc.,<sup>†</sup>  
AND JOOS V. LEBESQUE, M.D., PH.D.\*

\*Radiotherapy Department, The Netherlands Cancer Institute (Antoni van Leeuwenhoek Huis), Amsterdam, The Netherlands;

<sup>†</sup>Dr. Daniel den Hoed Cancer Center, Rotterdam, The Netherlands; and <sup>‡</sup>Dr. Bernard Verbeeten Institute, Tilburg, The Netherlands

**Purpose:** To investigate in three institutions, The Netherlands Cancer Institute (Antoni van Leeuwenhoek Huis [AvL]), Dr. Daniel den Hoed Cancer Center (DDHC), and Dr. Bernard Verbeeten Institute (BVI), how much the patient setup accuracy for irradiation of prostate cancer can be improved by an offline setup verification and correction procedure, using portal imaging.

**Methods and Materials:** The verification procedure consisted of two stages. During the first stage, setup deviations were measured during a number ( $N_{\max}$ ) of consecutive initial treatment sessions. The length of the average three dimensional (3D) setup deviation vector was compared with an action level for corrections, which shrunk with the number of setup measurements. After a correction was applied,  $N_{\max}$  measurements had to be performed again. Each institution chose different values for the initial action level (6, 9, and 10 mm) and  $N_{\max}$  (2 and 4). The choice of these parameters was based on a simulation of the procedure, using as input preestimated values of random and systematic deviations in each institution. During the second stage of the procedure, with weekly setup measurements, the AvL used a different criterion (“outlier detection”) for corrective actions than the DDHC and the BVI (“sliding average”). After each correction the first stage of the procedure was restarted. The procedure was tested for 151 patients (62 in AvL, 47 in DDHC, and 42 in BVI) treated for prostate carcinoma. Treatment techniques and portal image acquisition and analysis were different in each institution.

**Results:** The actual distributions of random and systematic deviations without corrections were estimated by eliminating the effect of the corrections. The percentage of mean (systematic) 3D deviations larger than 5 mm was 26% for the AvL and the DDHC, and 36% for the BVI. The setup accuracy after application of the procedure was considerably improved (percentage of mean 3D deviations larger than 5 mm was 1.6% in the AvL and 0% in the DDHC and BVI), in agreement with the results of the simulation. The number of corrections (about 0.7 on the average per patient) was not larger than predicted.

**Conclusion:** The verification procedure appeared to be feasible in the three institutions and enabled a significant reduction of mean 3D setup deviations. The computer simulation of the procedure proved to be a useful tool, because it enabled an accurate prediction of the setup accuracy and the required number of corrections.

Portal imaging, Setup accuracy, Decision rules, Prostate cancer.

### INTRODUCTION

Electronic portal imaging devices (EPIDs) are increasingly used in clinical practice to improve setup accuracy. In comparison with portal films, acquisition is faster and the evaluation is greatly facilitated by using computerized analysis methods (10, 11, 23).

Portal image analysis has shown that both the systematic and the random component of the setup deviation during a treatment session are important (3, 12, 13, 19). In principle, both components can be corrected by using online setup verification procedures with decisions on corrective actions taken during each treatment session. Recently, results of various online setup veri-

Reprint requests to: J. V. Lebesque, Radiotherapy Department, The Netherlands Cancer Institute (Antoni van Leeuwenhoek Huis) Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands. E-mail: jlebes@nki.nl.

Acknowledgements—The authors would like to thank the tech-

nicians who assisted in the implementation of the verification procedure. This study was supported by the Dutch Cancer Society, NKB Grants NKI 91-01 and DDHK 92-86.

Accepted for publication 4 December 1995.

fication procedures were reported (5, 8, 9). These reports showed that online procedures would involve a considerable workload.

Online procedures are possibly not required when random deviations are small. In this case, an offline procedure can be used, to correct systematic deviations only. Random deviations can be incorporated as a small margin around the clinical target volume, with the magnitude of the margin depending on the standard deviation (SD) of the distribution of random deviations (22). A difficulty with offline procedures is, however, that each measured setup deviation is the sum of a random and the systematic deviation. When and how corrective actions should be performed is not obvious and criteria for performing corrections are still a subject of discussion (14, 16, 20, 21). Recently, some theoretical procedures were proposed to deal with this problem (2, 6, 21, 28), but until now there have been no reports available about the clinical results of one of these procedures.

In an earlier paper (2) we described an offline procedure, using a shrinking action level, which could be adapted to various clinical conditions. This procedure aimed at reducing systematic deviations as much as possible. A computer simulation of the procedure was performed, in which parameters of the procedure were varied. To enable the simulation, a preestimate of the distribution of systematic and random deviations was required. As output, the resulting setup accuracy was provided as a function of the required number of portal images and setup corrections. It turned out that, for a specific level of accuracy, a compromise had to be made between the number of portal images and corrections. Hence, different institutions could make different choices for the parameters of the procedure depending on their specific demands.

The aim of this study was to investigate in three different institutions whether an offline three dimensional (3D) setup verification procedure is feasible in clinical practice, whether the procedure is efficient in reducing systematic deviations, and whether the improvement of patient setup accuracy could be predicted by a simulation of the setup verification procedure (2), using preestimated distributions of random and mean (systematic) deviations.

## METHODS AND MATERIALS

Three institutions in the Netherlands were involved in this study, namely The Netherlands Cancer Institute (Antoni van Leeuwenhoek Huis) in Amsterdam (AvL), the Dr. Daniel Den Hoed Cancer Center in Rotterdam (DDHC) and the Dr. Bernard Verbeeten Institute in Tilburg (BVI). Parameters will be labeled with AvL, DDHC, or BVI for the corresponding institution. The study started in January, 1993 in the AvL, in December, 1993 in DDHC, and in January, 1994 in the BVI, and ended in November, 1994.

### *Patients and treatment*

The patients in this study were treated for prostate carcinoma. The total number of patients involved in the evaluation of the verification procedure was 151 with 62 patients in the AvL, 47 patients in the DDHC, and 42 in the BVI.

Treatment preparation was similar in the three institutions. A computed tomography (CT) scan of the pelvic region was performed and the patient's skin was tattooed (marked in the DDHC) to indicate a tentative isocenter. During treatment planning, a new isocenter was defined by a 3D shift with respect to the tentative isocenter, when necessary. In the BVI, the vertical position of the new isocenter was defined with respect to the treatment room coordinate system, taking into consideration the estimated bending of the treatment table.

During the subsequent treatment simulation, lines were drawn on the skin to mark the definitive isocenter, in the AvL and the DDHC. In the BVI, asymmetric collimator settings and/or table shifts with respect to the tattooed tentative isocenter were used for correct positioning.

On the treatment machine, patients were positioned by aligning the lines, drawn during the simulation, with laser beams in the three main directions, in the AvL and the DDHC. In the BVI, the tattoos (indicating the tentative isocenter) were aligned with laser beams in the horizontal plane only. Subsequently, the table height was adapted, using the calculated position of the isocenter with respect to the table, thus defining the origin for the subsequent procedure. In some patients, a change of the tentative isocenter was required in the horizontal plane; in these cases, a prescribed translation of the table was performed, relative to this origin.

The treatment techniques were different for the three institutions (Table 1). In the AvL, the conformal simultaneous boost technique (15) was applied with three orthogonal fields (8 MV). In the DDHC, the gantry angle for the left lateral field ranged between 90° and 105° and for the right lateral field between 255° and 270°. In this institution, about half the number of patients were treated with rectangular fields (the majority with 25 MV), the other half treated with conformal fields (25 MV), defined by a multileaf collimator. In the BVI, a four-field box technique with rectangular fields was applied (10 MV). In the lateral fields, the collimator was rotated (collimator angle usually between 0° and 30°) to obtain a maximal sparing of the rectum.

### *Portal image acquisition and analysis*

In total, 4497 portal images were analyzed (2019 in the AvL, 1462 in the DDHC, and 1016 in the BVI). In the AvL and the DDHC, all fields of a treatment session were imaged. In the BVI, portal images were made of the anterior-posterior (AP) and left lateral field only. The three institutions had different types of electronic portal imaging devices (Table 1) and used different methods of image analysis.

Table 1. The treatment techniques, imaged fields and imaging systems of the three institutions

Institution	AvL	DDHC	BVI
<i>Treatment technique</i>			
Fields	Three fields (AP and two lateral)	Three fields (AP and two lateral oblique)	Four fields (AP, PA and two lateral)
Wedges	Lateral fields	Lateral fields	Not used
Number fractions	35	33	35
<i>Portal images</i>			
Imaged fields	All fields	All fields	AP and left lateral
Imaging system	Varian Portal Vision	Philips SRI 100	General Electric Targetview

AvL: The Netherlands Cancer Institute (Antoni van Leeuwenhoek Huis); DDHC: Daniel den Hoed Cancer Center; BVI: Bernard Verbeeten Institute; AP: Anterior–posterior; PA: Posterior–anterior.

In the AvL, a portal imaging device<sup>1</sup> was used that was based on a matrix of ionization chambers (25). Digitally reconstructed radiographs (DRRs) were made during the planning and used as reference images (1). The DRRs and the corresponding field outlines were imported into the portal image analysis system (23). On this reference image, a drawing of the bony anatomy was made. In the portal image, the field edges and the bony anatomy were detected and matched to the reference drawings automatically (10, 11) and corrected manually when necessary. The accuracy of this method was estimated to be about 0.7 mm (1 SD) for the caudal–cranial direction in both the AP and lateral fields. For the left–right direction in the AP field and the ventro–dorsal direction in the lateral field, the accuracy was estimated to be 0.3 mm and 1.0 mm (1 SD), respectively (11).

In the DDHC a camera-based portal imaging device<sup>2</sup> (26) was applied. The digitized simulator film was used as reference. Anatomical landmarks were indicated on the reference image and on the portal image and a landmark match of both images was performed. This match method was estimated to be accurate within 1 mm (1 SD).

A camera-based portal imaging system<sup>3</sup> (17) was also used in the BVI. The digitized simulator film served as reference image. The procedure for the analysis of portal images was very similar to the procedure used in the AvL; the only difference was that the anatomy in the portal image was delineated and matched by entirely manual techniques. The method was estimated to be equally accurate as in the AvL, with the exception of measurements in the lateral fields. Due to the larger interobserver variability (almost all of the technicians performed the anatomy alignment), the accuracy of these measurements was estimated to be within 1 mm (1 SD) for both directions.

### 3D setup deviations

In a previous study, it was shown that for prostate treatments, deviations in a particular direction, as mea-

sured in more than one field of a treatment session, were sufficiently correlated to allow the calculation of one single 3D setup deviation vector (3). By computing the average deviation in a particular direction, the effect of the measurement error is reduced. The three main directions are denoted by X, Y, and Z, with the X axis corresponding to the left–right direction, the Y axis corresponding to the caudal–cranial direction, and the Z axis corresponding to the dorsal–ventral direction.

For the AvL, images were acquired for the three treatment fields: the AP, left lateral (LL), and right lateral (RL) field. For the X component of the 3D setup deviation vector, only the deviation from the AP field was available,  $X_{AP}$ . The deviation in the Y direction was taken to be equal to the average of the Y deviations in the three fields,  $(Y_{AP} + Y_{RL} + Y_{LL})/3$ . The deviation in the Z direction was taken from the deviations along this axis of the two lateral fields,  $(Z_{RL} + Z_{LL})/2$ . For the DDHC, where three fields were also imaged, the same procedures were applied for the X and Y direction. The deviation in the Z direction was computed as  $([Z \sin(-\gamma)]_{RL} + [Z \sin \gamma]_{LL})/2$ , with  $\gamma$  being equal to the gantry angle of the lateral fields.

In the BVI, four fields were given but only portal images of the AP and the left lateral fields were acquired. The average of the deviations in the Y direction was calculated as  $(2Y_{AP} + Y_{RL})/3$ , thus taking the interobserver variability into account because the variability in the AP field was smaller than in the lateral field. The deviations along the X and Z directions were inferred from the AP and right lateral field, respectively, without calculation of an average.

### First stage of the verification procedure

The first stage of the verification procedure was a 3D extension of the one dimensional (1D) verification procedure with a shrinking action level described previously (1). Only translational setup errors were considered. Ro-

<sup>1</sup> Varian Portal Vision, Varian Radiation Division, Palo Alto, CA.

<sup>2</sup> Philips SRI 100, Philips Medical Systems Radiotherapy, Crawley, West Sussex, UK.

<sup>3</sup> General Electric Targetview, General Electric, Renelux.

tational errors were not considered in the procedure, because it has been shown in earlier studies that they are small for prostate treatments (3, 24).

At the beginning of the treatment, during  $N_{\max}$  consecutive treatment sessions, portal images were obtained. Of each treatment session, the 3D setup deviation was calculated (a *measurement*). After each measurement, the length of the average vector of the 3D setup deviation vectors over the first  $N$  measurements,  $|\vec{d}_{\text{avg}}| = |(\bar{x}, \bar{y}, \bar{z})|$ , was calculated. The length of this average vector was tested against an action level  $\alpha/\sqrt{N}$ , where  $\alpha$  is the parameter indicating the initial action level. When  $|\vec{d}_{\text{avg}}|$  was larger than the action level, a setup correction was performed (i.e., the patient was repositioned during the next treatment sessions by moving the isocenter over the average deviation in the three directions). After each correction, the calculation of the average vector was restarted and  $N_{\max}$  measurements had to be performed again. When a component of the required correction was smaller than a specific value, the correction was not applied in that direction. This value depended on the method of setup correction and was equal to 3 mm for the AvL and 2 mm for the DDHC and the BVI.

To determine for each institution values for the parameters  $N_{\max}$  and  $\alpha$ , a computer simulation of the procedure was performed, simulating the treatment of 10,000 patients. The simulation required as input a preestimate of the standard deviation  $\sigma_{\text{pre}}$  of the distribution of random deviations. Random deviations were assumed to be equal for different patients (3). A preestimate of the distribution of systematic deviations without corrections, with a standard deviation  $\Sigma_{\text{pre}}$ , was also needed to determine the expected workload. These data were obtained from earlier portal imaging studies in the three institutions (Table 2) (3, 4, 27). Both the systematic and random deviations were assumed to be normally distributed in the three directions. No correlations were assumed to exist between deviations in the three directions and the overall means were assumed to be equal to zero.

For each patient in the simulation, the vector length of

the mean of the three directions, the *mean 3D deviation*, was calculated. All setup deviations were taken into account in this mean 3D deviation, irrespective of setup corrections. For each institution, the percentage of mean 3D deviations larger than 5 mm was plotted against the average number of measurements and corrections per patient (Fig. 1) for varying values of  $\alpha$  and  $N_{\max}$ . All institutions aimed at a reduction of mean 3D deviations larger than 5 mm to about 5% or less. To obtain this level of setup accuracy, a compromise had to be made between the number of measurements and corrections. The graphs of Fig. 1 enabled each institution to make a choice for the parameters that was optimal for their situation.

As can be seen in Fig. 1, the desired level of setup accuracy was preestimated for the AvL with the parameter values  $N_{\max}^{\text{AvL}} = 2$  and  $\alpha_{\text{AvL}} = 6$  mm, resulting in 1.2% of the mean 3D deviations larger than 5 mm (Table 3). In the DDHC and the BVI, the setup accuracy without corrections was estimated to be lower than in the AvL and more measurements had to be performed to obtain the desired level of setup accuracy. Parameter values  $N_{\max}^{\text{DDHC}} = N_{\max}^{\text{BVI}} = 4$  with initial action levels  $\alpha_{\text{DDHC}} = 9$  mm and  $\alpha_{\text{BVI}} = 10$  mm were chosen in the DDHC and the BVI, respectively, resulting in 2.4 and 5.6% of the mean 3D deviations larger than 5 mm (Fig. 1 and Table 3), respectively. For the three institutions, the number of measurements and corrections (Fig. 1 and Table 3) were acceptably low.

#### Second stage of the verification procedure

When during  $N_{\max}$  consecutive treatment sessions no setup corrections were required, the first stage of the verification procedure was finished (1). This first stage was extended with a second stage to deal with large setup deviations that may arise during the course of the treatment. The cause of these deviations is yet unknown, but they can possibly be described as drifts, linear in time (7). Two different strategies for the second stage of the verification procedure were adopted. Both meth-

Table 2. Preestimates of the standard deviation of the distribution of random ( $\sigma_{\text{pre}}$ ) and mean (systematic) ( $\Sigma_{\text{pre}}$ ) deviations in each institution, which were used as input for the computer simulation of the verification procedure (Fig. 1)

Institution	AvL	DDHC	BVI
<i>Pre-estimated distributions</i>			
Authors	Bijhold <i>et al.</i> (3)	Creutzberg <i>et al.</i> (4)	Vos <i>et al.</i> (27)
$\sigma_{\text{pre}}$ (mm)	2	2	2.5
$\Sigma_{\text{pre}}$ (mm)	2	3	3
<i>Verification procedure</i>			
<i>First stage</i>			
$\alpha$ (mm)	6	9	10
$N_{\max}$	2	4	4
<i>Second stage</i>			
(weekly)	Outlier detection	Sliding average	Sliding average

The standard deviations were assumed to be equal in the three directions. The overall mean was assumed to be zero for all distributions. The parameter values for the first stage of the procedure are the initial action level  $\alpha$  and the number of measurements that had to be performed at the start of the treatment or after a setup correction,  $N_{\max}$ . During the second stage of the verification procedure, after finishing the first stage, different decision criteria were used in the three institutions.

ods were based on a weekly check after the first stage of the procedure.

In the AvL, it was tested after each weekly measurement if the vector length was outside the initial action level ( $\alpha_{AvL} = 6$  mm) and if one of the components of the vector was more than 5 mm different from the last calculated average. If both conditions were fulfilled, an “outlier” was detected and the calculation of  $|\bar{d}_{avg}|$  was restarted. The next day, a new measurement was performed and the average of the last two measurements was tested against the corresponding action level ( $\alpha_{AvL}/\sqrt{2} = 4.2$  mm).

For the DDHC and the BVI, a sliding average method was used (21, 27). After each weekly measurement, the vector length of the average of the last four measurements was calculated and tested against the lowest action level (i.e., 4.5 mm and 5 mm, respectively). Corrections were applied if this action level was exceeded. For all three institutions, the first stage of the verification procedure was restarted when a correction was required, otherwise the procedure continued with weekly measurements.

Both methods were also simulated with the corresponding preestimated data of the institutions as input and the already chosen parameter combinations for  $\alpha$  and  $N_{max}$ . Due to the lack of a model for gradual trends or sudden shifts in the setup during treatment, these errors could not be taken into account in the simulation. However, during the second stage of the verification procedure, random variations and remaining systematic deviations resulted in additional corrections in the simulation without occurrence of “real” gradual trends or sudden shifts of the setup. The simulation showed that these corrections resulted in a further substantial reduction of large mean 3D deviations (Table 3) compared with the first stage of the procedure.

#### *Implementation of the Procedure*

In the three institutions, portal image acquisition and analysis were performed by the technicians. In the AvL and the DDHC, the portal images were analyzed usually a few hours after the treatment session. In the BVI, the analysis was performed during the irradiation with the AP and RL fields, from which no portal images were acquired. The time required for analysis of the portal images of a treatment session was about 2–10 min. After matching the portal images with the reference images, the computer displayed decisions on corrective actions. In the AvL, the responsible physician checked and approved each setup correction. In the other two institutions, the DDHC and the BVI, the technicians applied corrections of the patient setup, without consultation of a physician.

Setup corrections were implemented differently in the three institutions. In the AvL, new lines were drawn on the patient’s skin and patients were aligned with these new lines. In the DDHC, no new lines were drawn. When a correction was required, the patient was positioned using the original lines drawn during simulation and subsequently, the table was shifted according to the mag-

nitude of the required correction. In the BVI, table shifts were already performed rather routinely during the treatment as a result of a change of the tentative isocenter during the planning. The overall correction was performed by moving the table over a distance equal to the sum of the setup correction and the already prescribed table movement. The setup correction procedure in the AvL resulted in a slight increase in the time (2–4 min) of that treatment session. In the two other institutions, the time of a treatment session was hardly prolonged due to setup corrections.

#### *Analysis*

An analysis of the data was performed to study the effect of the actual effectiveness of the verification procedure. The actual distribution of mean (systematic) deviations without corrections and the actual distributions of mean deviations after the first and second stage of the procedure were calculated. These last two distributions and the number of corrections and measurements were compared with the results of the simulation (using preestimated distributions) to evaluate the value of the simulation itself.

*Structure of setup deviations without the verification procedure.* To estimate the actual effect of the setup verification procedure, the setup deviations without the application of the verification procedure were estimated. The actual distribution of mean (systematic) deviations without corrections (SD denoted by  $\Sigma_{act}$ ) was estimated for each direction by eliminating the effect of the setup corrections, both in the first and the second stage of the procedure. After each correction, the magnitude of that correction was added to the subsequent deviations. This distribution was tested for being normal, because a normal distribution was used in the simulation.

For a patient, for whom the setup was corrected, the variances of the distribution of random deviations before and after each correction were calculated for each direction. Subsequently, the weighted average of these variances was calculated for each direction, resulting in the variance of the distribution of random deviations for this patient,  $\sigma_i^2$ . The variance of the distribution of random deviations for patients for which no corrections were applied, was defined in the usual way,  $\sigma_i^2 = \sum_{j=1}^N (x_j - \bar{x})^2 / (N - 1)$ , with  $x_j$  a measured setup deviation in one of the three directions. The variance of the actual distribution of random deviations of the whole patient group,  $\sigma_{act}^2$ , was computed by subsequently taking the average of the variances of all patients. In the simulation, the preestimates of the distributions of random deviations were assumed to be normal with an SD that was assumed to be equal for each patient. Both assumptions were tested for all patients.

In the simulation, no time-dependent gradual trends or sudden shifts of the setup during treatment were assumed. The occurrence of trends was studied by performing a linear fit of the deviations in three directions with time. Each slope was tested for being significantly different

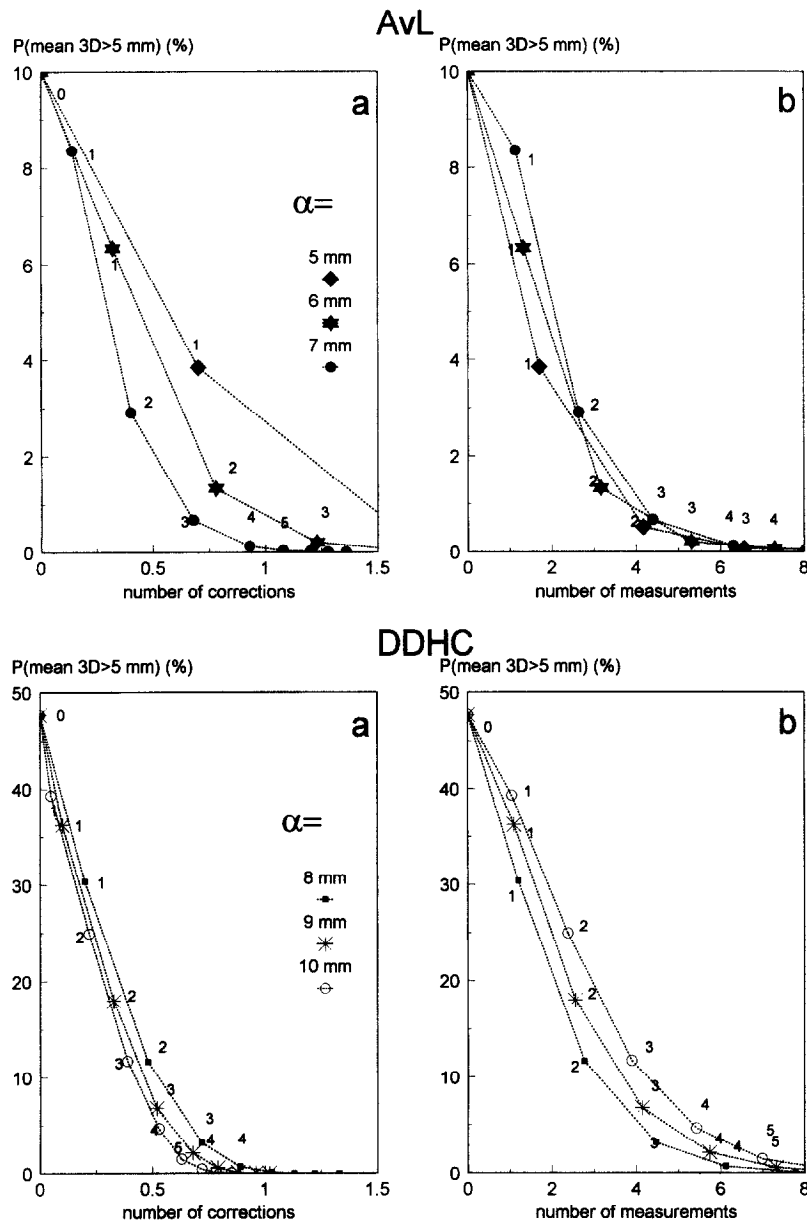


Fig. 1. Simulation of the first stage of the procedure. The setup accuracy, expressed as the percentage of mean 3D deviations larger than 5 mm,  $P(\text{mean } 3D > 5 \text{ mm})$ , is plotted against the average number of corrections (a) and measurements (b) per patient. The initial action level  $\alpha$  and the number of measurements,  $N_{\text{max}}$ , that had to be performed at the start of the treatment or after each correction, were varied. The value of  $N_{\text{max}}$  is indicated near each point in the graph.

from zero. A detailed analysis of trends is presented in the appendix.

**First stage of the verification procedure.** The evaluation of the first stage of the procedure was confined to a test of the validity of the computer simulation, which served as a basis for the choice of the procedure parameters (Fig. 1). The effects of corrections during the second stage of the procedure were eliminated. The setup accuracy after application of the first stage of the procedure was subsequently determined by computing for each patient the mean 3D setup deviation. The resulting actual distribution of mean 3D deviations after the application of the first stage of the procedure was tested for being significantly different from the corresponding preesti-

mated distributions. Without corrections, these deviations are distributed as a  $\chi^2$  distribution and after the application of the procedure the tail of the distribution is cut off. In addition, the average number of measurements and corrections per patient during the first stage was determined and compared with the results of the simulation.

**Total verification procedure.** The total procedure, consisting of the first and second stage, was analyzed along the same lines as the first stage of the procedure. The actual distribution of mean 3D deviations and the average number of measurements and corrections per patient after application of the total procedure were calculated and compared with the results of the simulation.

An individual patient setup with an SD of the distribu-

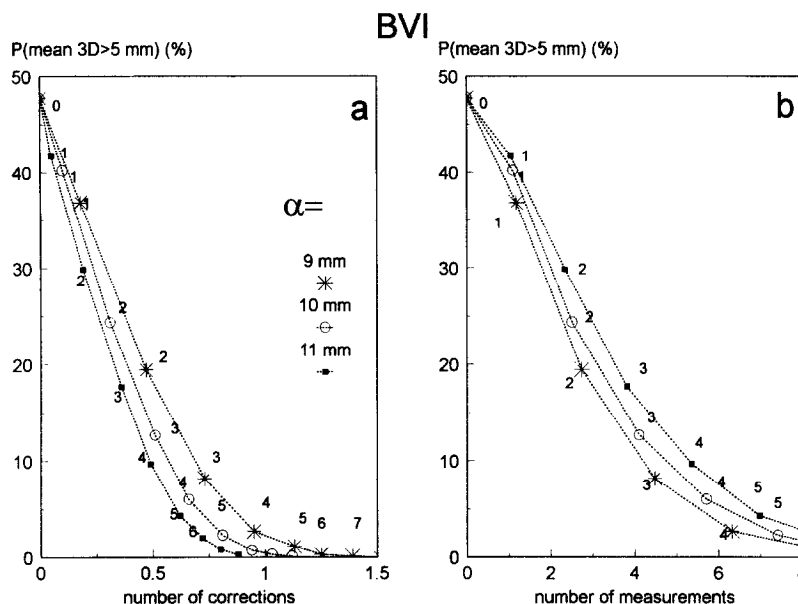


Fig. 1. (Cont'd)

tion of random deviations ( $\alpha_i$ ), which is larger than the standard deviation of the total group ( $\sigma_{act}$ ) can lead to repeated setup corrections for this patient (1). Therefore, it was tested whether setups of individual patients with relatively large or small random deviations were significantly more or less corrected than the setups of the total group.

*Statistical procedures.* Several statistical tests were used to determine the significance of the difference between various quantities. The significance of the difference between the actual distribution of mean 3D deviations and the corresponding preestimated distribution, as derived from the simulation, was calculated by using the

Table 3. The preestimated setup accuracy without corrections and the results of the simulation of both the first stage of the verification procedure and the total verification procedure

Institution	AvL	DDHC	BVI
<i>Number of mean 3D deviations &gt;5 mm (%)</i>			
Without corrections	10.0	42.7	42.7
First stage of procedure	1.2	2.4	5.6
Total procedure	0.9	0.5	0.9
<i>Average number of measurements per patient</i>			
First stage of procedure	3.2	5.7	5.8
Total procedure	10.7	12.1	13.4
<i>Average number of corrections per patient</i>			
First stage of procedure	0.8	0.7	0.7
Total procedure	1.1	0.8	0.8

As input for this simulation, the preestimated distributions of random and mean (systematic) deviations were used (Table 2). As a measure of the setup accuracy, the percentage of mean 3D deviations larger than 5 mm was used. The resulting workload was quantified by the average number of measurements and corrections per patient.

Kolmogorov–Smirnov (K–S) test (18). An *F*-test was used to determine whether the SD of the distribution of the random deviations of individual patients was significantly different from the SD of the actual distribution of the whole group.

A Monte Carlo simulation (18) was used to determine confidence levels of various quantities. This method is especially useful when a quantity is calculated from a distribution of measurements that is not normal, as with, for instance, the average of all setup deviations of all patients after application of corrections. The Monte Carlo simulation consists of the creation of an artificial “result,” generated by randomly choosing data from the distribution of measurements. This procedure is repeated (10,000 times), resulting in a frequency distribution of the calculated quantity. Subsequently, the confidence levels of the calculated quantity can be determined.

This Monte Carlo simulation was used to determine whether the overall means (the average overall deviations of all patients) were significantly different from zero and to test whether the preestimated and actual percentage of mean 3D deviations larger than 5 mm were significantly different. A Monte Carlo simulation was performed on the patient data to test the significance of differences between the average number of measurements and corrections per patient according to the actual patient data and the simulation. Finally, the method was used to test whether the slopes, after fitting the deviations with time (see appendix), were significantly different from zero.

## RESULTS

To enable evaluation of effectiveness of the verification procedure, results after elimination of the effect of corrections are presented first. Using these results, the validity of assumptions of of setup deviations statistics, made for

the computer simulation (Table 2) can also be tested. Subsequently, we present the evaluation of effectiveness of the first stage of the verification procedure in improving the setup accuracy, and a comparison with the outcomes of the simulation. Finally, such an evaluation and comparison are given for the total procedure, consisting of the first and second stages.

#### Structure of setup deviations without the verification procedure

The actual distributions of the mean (systematic) deviations (with SD  $\Sigma_{act}$ ) without corrections were estimated (Table 4) by elimination of the effect of corrections. Only for the AvL, along the dorsal-ventral axis, was this distribution significantly ( $p = 0.02$ ) different from a normal distribution. For none of the institutions were the differences between  $\Sigma_{act}$  and the corresponding preestimated values  $\Sigma_{pre}$  (Table 2) significant.

For the AvL and the BVI, the overall mean (i.e., the deviations averaged over all treatment setups of all patients), was significantly ( $p < 0.001$ ) different from zero for the caudal-cranial and the dorsal-ventral direction (Table 4). For the DDHC, the overall mean in the left-right direction was significantly ( $p < 0.001$ ) different from zero. These overall means unequal to zero were most probably due to small systematic treatment machine inaccuracies.

As a result of the overall mean being unequal to zero and because of the differences between the preestimated and actual distribution of mean (systematic) deviations without corrections, the corresponding distributions of

mean 3D deviations (Fig. 2) were significantly ( $p < 0.01$ ) different for the AvL and the DDHC. The differences at the level of 5 mm (Tables 3 and 5) were also substantial for the three institutions.

The SDs of the actual distribution of random deviations,  $\sigma_{act}$  (Table 4), for the three institutions were only slightly different from the SDs that were assumed in the simulation (Table 2). There were no patients with an actual distribution of random deviations significantly different from a normal distribution. It was tested whether the SD of the distribution of random deviations of individual patients ( $\sigma_i$ ) along any direction was significantly ( $p < 0.05$ ) different from  $\sigma_{act}$  (Table 4). For patients with  $\sigma_i$  significantly larger than  $\sigma_{act}$ , this SD was equal to 3.6 mm, 3.8 mm, and 4.4 mm for the AvL, the DDHC, and the BVI, respectively. One patient (in the AvL) had a value of  $\sigma_i$  that was significantly larger than  $\sigma_{act}$  in more than one direction. There were no patients with  $\sigma_i$  significantly larger than  $\sigma_{act}$  in one direction and  $\sigma_i$  significantly smaller than  $\sigma_{act}$  in another direction.

Trends in time were analyzed by performing linear fits of the deviations, after the elimination of the effect of corrections. In about half the number of patients, a significant trend in time was observed (appendix Table A1). For about 20% of the patients in the AvL and the DDHC and 5% of the patients in the BVI, these trends would have resulted in a drift of the length of the mean (systematic) 3D deviation vector of more than 5 mm during the course of the treatment.

#### First stage of the verification procedure

To test the validity of the simulation (using as input the preestimated distributions, Table 2), the actual and preestimated distributions of mean 3D deviations after application of the first stage of the procedure were compared. No significant differences between these distributions were found for any institution. However, at a particular level of the distributions (i.e., the probability of mean 3D deviations larger than 5 mm), a significant difference was found for the AvL between the actual value (11.3%, Table 5) and the preestimated value (1.2%, Table 3). This was probably caused by the underestimation of the actual number of mean (systematic) 3D deviations larger than 5 mm (Table 4) without corrections with respect to the corresponding preestimated value (Table 2).

The workload of the first stage of the procedure was evaluated (Table 5). In the AvL, the average number of measurements and corrections per patient was significantly smaller than that estimated from the simulation (Table 3). For the two other institutions, the average number of measurements and corrections per patient was also lower than predicted, but only for the BVI the average number of corrections per patient was significantly smaller.

#### Total verification procedure

For each institution, the actual distribution of mean 3D deviations, after application of both the first and the second stage of the procedure, was calculated (Fig. 2, Table 5).

Table 4. The actual distribution of mean (systematic) deviations without corrections, calculated by the elimination of the effect of the corrections, with the overall means and standard deviations (SDs)

Institution	AvL	DDHC	BVI
Number of patients	62	47	42
Number of mean 3D deviations > 5 mm (%)			
Without corrections	25.8	25.5	35.7
First stage of procedure	11.3	2.1	4.8
Total procedure	1.6	0	0
Average number of measurements per patient			
First stage of procedure	2.6	5.4	5.2
Total procedure	10.9	10.4	12.1
Average number of corrections per patient			
First stage of procedure	0.4	0.5	0.4
Total procedure	0.6	0.8	0.7
Total procedure, subgroup with:			
$\sigma_i$ significantly smaller than $\sigma_{act}$	0.4	1.1	0.6
$\sigma_i$ significantly larger than $\sigma_{act}$	1.4	1.5	1.5

The standard deviation of the actual distribution of random deviations ( $\sigma_{act}$ ) for all patients was calculated and subsequently, the percentage of patients with a standard deviation  $\sigma_i$  of the individual distribution of random deviations significantly ( $p < 0.05$ ) larger or smaller than  $\sigma_{act}$  was determined.



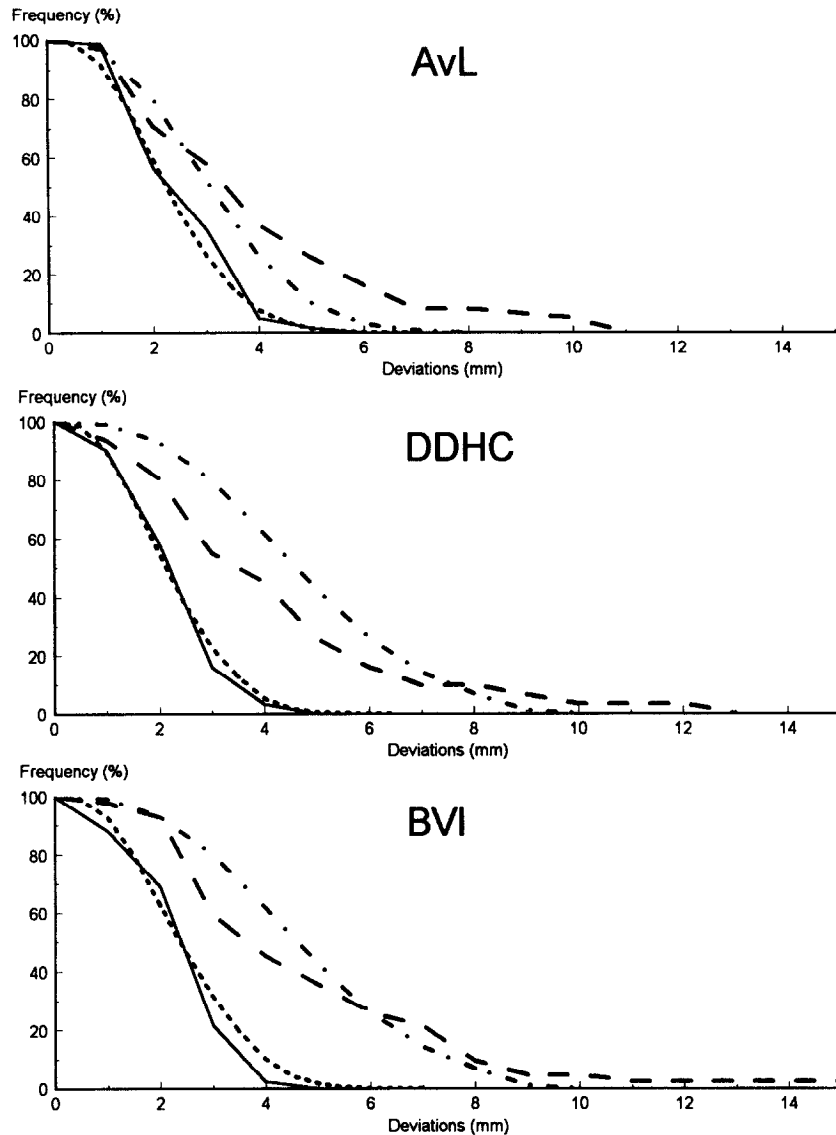


Fig. 2. Cumulative distributions of mean 3D deviations. Along the vertical axis, the percentage of deviations larger than a specific value is plotted against that value on the horizontal axis. Plotted are the preestimated distributions without corrections (dashed/dotted line) and the preestimated distributions after application of the total verification procedure (dotted line). The actual distributions without corrections and after application of the procedure are given by the dashed and solid line, respectively.

For the three institutions there was a major improvement with respect to the actual and preestimated distribution of mean (systematic) 3D deviations without corrections. The actual distributions of mean 3D deviations after application of the total procedure did not differ significantly from the corresponding preestimated distributions. For the three institutions, the actual distributions of mean 3D deviations were similar. In comparison with the first stage of the procedure, the percentage of mean 3D deviations larger than 5 mm was further reduced. In the AvL, 1.6% of the mean 3D deviations was larger than 5 mm, and for the DDHC and the BVI there were no more mean 3D deviations larger than 5 mm. These results were in agreement with the level of setup accuracy as preestimated by the simulation (Table 3).

The additional average number of corrections per patient during the second stage of the procedure (Table 5) agreed

with the simulation (Table 3) for the three institutions. For the AvL, the total average number of measurements per patient during the first and the second stage of the procedure agreed with the simulation. In the DDHC and the BVI, the total average number of measurements per patient was significantly smaller than was predicted. In the DDHC, this was merely because the weekly check was not always performed consistently.

It was tested whether the number of corrections of the subgroups of patients with relatively small or large standard deviations of the distribution of random deviations ( $\sigma_i$ ) was different from the number of corrections of the total group. The average number of corrections per patient in the subgroup with a small  $\sigma_i$  (significantly smaller than the SD of the distribution of random deviations ( $\sigma_{act}$ ) of the whole group) was not significantly

Table 5. The actual setup accuracy without corrections, after elimination of the effect of the corrections, and the results of the first stage of the verification procedure and the total verification procedure, consisting of the first and second stage

Institution	AvL	DDHC	BVI
Overall mean (mm)			
Left-right	0.3	-0.7	-0.3
Caudal-cranial	-0.9	-0.3	1.6
Dorsal-ventral	0.9	-0.4	-1.3
SD systematic ( $\Sigma_{act}$ ) (mm)			
Left-right	2.4	2.9	2.2
Caudal-cranial	2.4	2.1	2.7
Dorsal-ventral	2.6	3.3	3.8
SD random ( $\sigma_{act}$ ) (mm)			
Left-right	2.0	2.1	2.7
Caudal-cranial	1.8	2.1	2.1
Dorsal-ventral	1.7	2.1	2.3
Patients with $\sigma_i$ significantly smaller than $\sigma_{act}$ (%)			
Left-right	6	8	7
Caudal-cranial	10	6	2
Dorsal-ventral	5	4	2
For any direction	19	18	12
Patients with $\sigma_i$ significantly larger than $\sigma_{act}$ (%)			
Left-right	2	4	7
Caudal-cranial	6	4	2
Dorsal-ventral	5	4	0
For any direction	11	13	10

The setup accuracy was quantified by the percentage of mean 3D deviations larger than 5 mm. The workload involved in the application of the procedure was quantified by the average number of measurements and corrections per patient. For the patients with a standard deviation of the individual distribution of random deviations  $\sigma_i$  in one or more of the directions significantly ( $p < 0.05$ ) smaller or larger than the standard deviation of the total group  $\sigma_{act}$ , the average number of corrections per patient was determined.

smaller than the average number of corrections per patient of the total group (Table 5). On the other hand, for the subgroup of patients with a large  $\sigma_i$  (significantly larger than  $\sigma_{act}$ ), more corrections were performed than in the total group (Table 5). For the three institutions together, this difference was significant, but for each particular institution, the difference could not prove to be significant.

The application of the procedure reduced the magnitudes of the overall means with respect to the calculated magnitudes of the overall means after the elimination of the effect of the corrections (Table 4). However, small overall means that were significantly different from zero were still present in the AvL in the ventral and the caudal direction with magnitudes equal to 0.6 mm and 0.5 mm, respectively. In the DDHC, none of the overall means were significantly different from zero. In the BVI there was a significant overall mean different from zero in the cranial direction, with a magnitude equal to 0.8 mm.

## DISCUSSION

The offline setup verification procedure resulted in a major improvement of the setup accuracy in the three

institutions involved in this study. Despite different treatment techniques, patient setup procedures, and portal imaging systems in the institutions, the same level of accuracy could be obtained by tailoring the procedure to the preestimated treatment accuracy without corrections in each institution, using a computer simulation.

By varying the parameters of the computer simulation procedure, each institution was able to balance the desired level of accuracy against the required number of portal images and setup corrections. The preestimated distributions of systematic and random deviations were not precisely equal to the actual distributions in the clinic. However, the resulting setup accuracy did not critically depend on such differences in input as discussed in an earlier report (2). The resulting setup accuracy was also not critically dependent on the occurrence of trends in time. For none of the institutions was the resulting setup accuracy significantly different from that expected.

In the three institutions, the procedure was mainly executed by technicians. Decisions on corrective actions were indicated by the computer. Technicians appreciated the availability of objective criteria, in contrast to old procedures in which it was up to the subjective interpretation of the responsible physician to decide on setup corrections. Due to the computerized acquisition and (partly) computerized image analysis, the implementation could be performed without a large increase in workload. Especially in the BVI, image analysis was performed in a time-efficient manner. The images were analyzed in the treatment control room after the acquisition of the portal images of two fields, while the patient was irradiated with the other fields from which no portal images were acquired.

In the simulation of the verification procedure it was assumed that the distributions of random deviations were equal for different patients. However, this study revealed that for about 11% of the patients the SD of the distribution of random deviations in one of the directions was larger than the SD of the whole group. Probably due to these large random deviations, the number of setup corrections for this subgroup was significantly larger than the number of setup corrections of the whole group. To deal with these large random deviations in the future, a new estimate of the SD of the random deviations of each individual patient ( $\sigma_i$ ) can be made after each new setup measurement. Subsequently, it can be tested whether  $\sigma_i$  is significantly larger than the expected value. Possibly unnecessary setup corrections can be prevented by adapting the parameters of the verification procedure. For these patients, the procedure can be continued with a larger number of setup measurements and a larger initial action level (2). In addition, the treatment plan can be adapted. A new treatment plan can be made for the remaining treatment sessions, with an adapted margin around the clinical target volume. Moreover, the dosimetric consequences of the first part of the treatment (with a wrong estimate of the margin) should also be considered during such an adaptation of the treatment plan.

A possible further improvement of setup accuracy can be obtained by “fine-tuning” the procedure. Because, in general, the random deviations are not equal in the three directions, the spherical action level can be replaced by an elliptical action level, with the main axis of the ellipsoid proportional to the variance of the distribution of random deviations in each direction. A further refinement can be obtained by performing corrections with a damping: overcorrection can be prevented as much as possible by performing a correction that is not equal to the average deviation but that is an optimal estimate of the systematic deviation (21).

Sudden shifts in the setup during the course of the treatment were not analyzed. Consequently, these errors were probably interpreted as trends in time. Although there is no clear explanation for these sudden shifts or gradual trends (7) in the setup, their occurrence clearly necessitates a verification procedure after the first stage of the procedure. Using the strategy in the second stage

of the procedure of the AvL, with an “outlier” detection, sudden shifts in the setup can be detected quickly. On the other hand, gradual trends are better handled by the “sliding average” procedure of the DDHC and the BVI. A promising, more general concept for a verification procedure during the second stage of the procedure can be derived from Yan *et al.* (28). In their approach, the setup deviation of the next treatment session was predicted. Corrective actions can be performed when this deviation is unacceptably large.

## CONCLUSIONS

The offline setup verification procedure resulted in an improvement of patient setup accuracy by a major reduction of large systematic deviations, under different clinical conditions. The procedure could be implemented in the clinic with an acceptably small increase in workload. The resulting setup accuracy and the required workload could well be predicted by means of a computer simulation.

## REFERENCES

1. Bel, A.; Bartelink, H.; Vijlbrief, R. E.; Lebesque, J. V. Transfer errors of planning CT to simulator: A possible source of setup inaccuracies? *Radiother. Oncol.* 31:176–180; 1994.
2. Bel, A.; van Herk, M.; Bartelink, H.; Lebesque, J. V. A verification procedure to improve patient setup accuracy using portal images. *Radiother. Oncol.* 29:253–260; 1993.
3. Bijhold, J.; Lebesque, J. V.; Hart, A. A. M.; Vijlbrief, R. E. Maximizing setup accuracy using portal images as applied to a conformal boost technique for prostatic cancer. *Radiother. Oncol.* 24:261–271; 1992.
4. Creutzberg, C. L.; Althof, V. G. M.; de Hoog, M.; Visser, A. G.; Huizenga, H.; Wijnmaalen, A. J.; Levendag, P. C. Consequences of variations in patient positioning for clinical practice. *Radiother. Oncol.* 32(Suppl. 1):117; 1994.
5. De Neve, W.; Van den Heuvel, F.; De Beukeleer, M.; Coghe, M.; Thon, L.; De Roover, P.; Van Lancker, M.; Storme, G. Routine clinical on-line portal imaging followed by immediate field adjustment using a tele-controlled patient couch. *Radiother. Oncol.* 24:45–54; 1992.
6. Denham, J. W.; Daily, M. J.; Hunter, K. Wheat, J.; Fahey, P. P.; Hamilton, C. S. Objective decision-making following a portal film. The results of a pilot study. *Int. J. Radiat. Oncol. Biol. Phys.* 26:869–876; 1993.
7. El-Gayed, A. H.; Bartelink, H.; Bel, A.; Vijlbrief R.; Lebesque, J. V. Evaluation of the time trend of setup deviations during the course of pelvic irradiation using an electronic portal imaging device. *Radiother. Oncol.* 24:45–54; 1992.
8. Ezz, A.; Munro, P.; Porter, A. T.; Battista, J.; Jaffray, D. A.; Fenster, A.; Osborne, S. Daily monitoring and correction of radiation field placement using a video-based portal imaging system: A pilot study. *Int. J. Radiat. Oncol. Biol. Phys.* 22:159–165; 1992.
9. Gildersleve, J.; Dearnaley D. P.; Evans P. M.; Law M.; Rawlings C.; Swindell, W. A randomised trial of patient repositioning during radiotherapy using a megavoltage imaging system. *Radiother. Oncol.* 31:161–168; 1994.
10. Gilhuijs, K. G. A.; Touw, A.; van Herk, M.; Vijlbrief, R. E. Optimization of automatic portal image analysis. *Med. Phys.* Submitted; 1995.
11. Gilhuijs, K. G. A.; van Herk, M. Automatic on-line inspection of patient setup from digital portal images. *Med. Phys.* 20:667–677; 1993.
12. Griffiths, S. E.; Khoury, G. G.; Eddy, A. Quality control of radiotherapy during pelvic irradiation. *Radiother. Oncol.* 20:203–209; 1991.
13. Huizenga, H.; Levendag, P. C.; De Porre, M. Z. R.; Visser, A. G. Accuracy in radiation field alignment in head and neck cancer: A prospective study. *Radiother. Oncol.* 11:181–188; 1988.
14. Lebesque, J. V.; Bel, A.; Bijhold, J.; Hart, A. A. M. Detection of systematic patient setup errors by portal film analysis. *Radiother. Oncol.* 23:198; 1992.
15. Lebesque, J. V.; Keus, R. B. The simultaneous boost technique: The concept of relative normalized total dose. *Radiother. Oncol.* 22:45–55; 1991.
16. Mitine, C.; Leunens, G.; Verstraete, J.; Blankaert, N.; Van Dam, J.; Dutreix, A.; Van der Schueren, E. Is it necessary to repeat quality control procedures for head and neck patients? *Radiother. Oncol.* 21:201–210; 1991.
17. Munro, P.; Rawlinson, J. A.; Fenster, A. A digital fluoroscopic imaging device for radiotherapy localization. *Int. J. Radiat. Oncol. Biol. Phys.* 18:641–649; 1990.
18. Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical recipes in C.* Cambridge: Cambridge University Press; 1991:471–565.
19. Rabinowitz, I.; Broomberg, J.; Goitein, M.; McCarthy, K.; Leong J. Accuracy of radiation field alignment in clinical practice. *Int. J. Radiat. Oncol. Biol. Phys.* 11:1857–1867; 1985.
20. Shalev, S. Progress in the evaluation of electronic portal imaging—Taking one step at a time. *Int. J. Radiat. Oncol. Biol. Phys.* 28:1043–1045; 1994.
21. Shalev, S.; Glutchev, G. Decision rules for the correction of field set-up parameters. *Med. Phys.* 21:952; 1994.
22. Urie, M. M.; Goitein, M.; Doppke, K.; Kutcher, J. G.; LoSasso, T.; Mohan, R.; Munzenrider, J. E.; Sontag, M.; Wong, J. W. The role of uncertainty analysis in treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.* 21:91–107; 1991.
23. van Herk, M.; Bel, A.; Gilhuijs, K. G. A.; Vijlbrief, R. E. A comprehensive system for the analysis of portal images. *Radiother. Oncol.* 29:221–229; 1993.
24. van Herk, M.; Bruce, A.; Kroes, G.; Shuman, T.; Touw, A.; Lebesque, J. V. Quantification of organ motion during conformal radiotherapy of the prostate by three-dimensional

- sional image correlation. *Int. J. Radiat. Oncol. Biol. Phys.* 33:1311–1320; 1995.
25. van Herk, M.; Meertens, H. A matrix ionization chamber imaging device for on-line patient setup verification during radiotherapy. *Radiother. Oncol.* 11:369–378; 1988.
26. Visser, A. G.; Huizenga, H.; Althof, V. G. M.; Swanenburg, B. N. Performance of a prototype fluoroscopic radiotherapy imaging system. *Int. J. Radiat. Oncol. Biol. Phys.* 18:43–50; 1990.

27. Vos, P. H.; Bel, A.; Poortmans, P. M. P.; Vlaun, V.; Rodrigues, P. T. R.; Lebesque, J. V. Set-up accuracy for pelvic fields after application of a 3-D correction procedure using digital portal imaging. *Int. J. Radiat. Oncol. Biol. Phys.* 30(Suppl. 1):229; 1994.
28. Yan, D.; Wong, J. W.; Gustafson, G.; Martinez, A. A new model for “accept or reject” strategies in on-line and off-line megavoltage treatment evaluation. *Int. J. Radiat. Oncol. Biol. Phys.* 31:943–952; 1995.

## APPENDIX: TRENDS

### Calculation of trends

The data after the elimination of the effect of corrections were used to determine the occurrence of trends. To study whether trends occurred at all, the deviations in each direction were linearly fitted with time, with the time expressed as the number of days since the start of treatment. It was tested whether slopes were significantly different from zero by means of a Monte Carlo simulation. The consequences of the trends were evaluated by calculating, for patients with a significant trend in one or more directions, the difference between the length of the 3D setup deviation vector at the start and at the end of the treatment (3D drift).

The distribution of slopes, intersections (or *initial systematic deviations*) and “time-corrected” random deviations were calculated for each direction. For the calculation of the distribution of slopes, all slopes were taken into account, regardless of whether the slope was significantly different from zero or not. The distribution of time-corrected random deviations was calculated by eliminating the trend (i.e., a factor equal to the slope multiplied by the time was subtracted from each setup deviation). The standard deviation of this distribution of time-corrected random deviations will be smaller than the corresponding standard deviation of the actual distribution of random deviations,  $\sigma_{act}$ , without trends, because trends lead to an overestimate of the random deviations.

### Results

In the three institutions, in about half the number of patients trends occurred in one or more directions, with a slope significantly ( $p < 0.05$ ) different from zero (appendix Table 1). In the AvL, most trends were along the left–right axis, while in the BVI most trends were along the dorsal–ventral axis. In the DDHC, the number of significant trends was about equally distributed over the three directions. In the AvL and the DDHC, in about 20% of the patients a trend in one of the directions resulted in a drift of the length of the 3D vector during course of treatment that was larger than 5 mm (appendix Table 1). In the BVI, only 5% of the patients had such a drift.

The distributions of slopes had standard deviations that

Table A1. Analysis of linear trends in time, with the number of significant ( $p < 0.05$ ) trends and the resulting drift of the length of the 3D vector for patients with such a trend

Institution	AvL	DDHC	BVI
<i>Number of significant (<math>p &lt; 0.05</math>) trends (%)</i>			
<i>Left–right</i>	31	19	19
<i>Caudal–cranial</i>	27	21	17
<i>Dorsal–ventral</i>	11	19	31
<i>Any direction</i>	53	52	52
<i>Drift during treatment &gt; 5 mm</i>	19	21	5
<i>SD all slopes (<math>10^{-2}</math> mm/day)</i>			
<i>Left–right</i>	7.1	12.1	7.5
<i>Caudal–cranial</i>	6.4	12.2	4.7
<i>Dorsal–ventral</i>	6.0	11.3	6.8
<i>SD initial systematic (mm)</i>			
<i>Left–right</i>	2.0	2.8	2.1
<i>Caudal–cranial</i>	2.0	2.6	2.6
<i>Dorsal–ventral</i>	2.5	2.8	3.6
<i>SD random (mm)</i>			
<i>Left–right</i>	1.9	2.1	2.6
<i>Caudal–cranial</i>	1.8	1.9	2.1
<i>Dorsal–ventral</i>	1.5	1.9	2.0

The standard deviation (SD) of the distribution of all slopes was calculated. The SD of the distribution of initial systematic and random deviations, was also calculated (comparable with the SD of the distribution of systematic and random deviations,  $\Sigma_{act}$  and  $\sigma_{act}$ , calculated without taking trends into account (Table 4)).

were on the average equal to 0.07 mm/day in the AvL and the BVI. In the DDHC, the SDs were slightly larger, about 0.12 mm/day. There were no preferences of the patients to move into a certain direction during the course of treatment, because none of the distributions of slopes had a mean that was significantly different from zero. The SDs of the distributions of initial systematic and time-corrected random deviations were slightly smaller than the SDs of the corresponding actual distributions of the systematic and random deviations,  $\sigma_{act}$  and  $\Sigma_{act}$ , without taking trends into account (Table 4). The overall means of the distribution of initial systematic deviations were only slightly (less than 0.2 mm) different from the overall means calculated without taking trends into account.