A verification procedure to improve patient set-up accuracy using portal images

A. Bel*, M. van Herk, H. Bartelink, J.V. Lebesque

Radiotherapy Department, The Netherlands Cancer Institute (Antoni van Leeuwenhoek Huis), Plesmanlaan 121, 1066 CX, Amsterdam, The Netherlands

(Received 5 February 1993; revision received 18 May 1993; accepted 29 May 1993)

Abstract

The purpose of this study was to establish which level of geometrical accuracy can be obtained during radiotherapy, using portal image analysis, with a minimum number of patient set-up measurements and corrections. A set-up verification and correction procedure using decision rules for improving the set-up of a patient during radiotherapy was investigated by means of a computer simulation. In this simulation study, set-up deviations were assumed to be the sum of random and systematic deviations and varying ratios of random and systematic deviations were studied. The distribution of random deviations (SD equal to \( \sigma \)) was assumed to be equal for all patients of a specific treatment site. Set-up deviations are measured during the first \( N \) consecutive fractions after the start of the treatment or after a patient set-up correction. A set-up is corrected when the deviation averaged over these measurements is larger than an \( N \)-dependent action level. This action level is specified by \( \alpha_{\text{initial action level}} \), in which \( \alpha \) is a variable initial action level parameter. After the start of the treatment or after each correction, \( N_{\text{max}} \) measurements are made to decide on a possible (further) correction. By varying \( \alpha \) and \( N_{\text{max}} \), the relation between the overall accuracy and the workload has been analyzed. It was possible to obtain a resulting overall accuracy level which is almost independent of the initial distribution of systematic deviations. When the standard deviation \( \Sigma \) of the distribution of systematic deviations ranges from \( \sigma \) to \( 3\sigma \), the parameters \( \alpha = 2\sigma \) and \( N_{\text{max}} = 2 \) result in an accuracy with less than 5% of the set-ups larger than \( 3\sigma \). Using these parameters, for \( \Sigma = 2\sigma \), 3.2 measurements and 0.9 corrections per patient on average are required. The probability that a patient has to be corrected more than once is 22.5%. A specific set-up with an initial large systematic deviation equal to 4.5\( \sigma \) results in an average deviation smaller than 2\( \sigma \) with a probability equal to 99%.

Key words: Radiation therapy; Portal imaging; Quality control; Patient set-up accuracy; Decision rules; Computer simulation

1. Introduction

Patient set-up deviations can be measured during radiotherapy by comparing portal images to simulator images. Important issues related to these observed set-up deviations which require attention are: what is the statistical character of the deviations, and which deviations are acceptable or need immediate attention? These questions become more important since the clinical introduction of digital portal imaging systems [1,10,12,15]. Furthermore, the analysis of a large number of portal images is facilitated with the introduction of automatic analysis procedures [6].

Set-up deviations can be described as a sum of systematic and random contributions. In several studies, the parameters of the distribution function of systematic and random deviations were determined by analysis of a large series of portal images [7,8,13]. However, these studies did not discuss criteria for set-up corrections.

In a recent study of Bijhold et al. [2] random and systematic distributions of set-up deviations of pelvic treatments were analyzed. Firstly, they showed that there were no significant differences between the distributions of random deviations of individual patients in this patient group. They also demonstrated that the distribution of random deviations was not significantly different from the distribution of systematic deviations over the whole group of patients.
Their results could be accounted for by their concept of systematic deviations. A systematic deviation is the difference between the intended set-up on the simulator and the actual set-up during the treatment. Day-to-day (random) variations can arise from the movement of the skin marks with respect to the bones. This random variation is also present on the simulator. Because the position of the skin marks on the simulator is a sample from this distribution of random deviations, the smallest distribution of systematic deviations that can be obtained is equal to the distribution of random deviations. From these arguments it was suggested that in their study there were no errors in the transfer of data from simulator to accelerator ('transfer errors'). If there would have been transfer errors, the systematic deviations would have been larger than the random deviations.

Bijhold et al. [2] proposed an off-line set-up verification procedure, with decisions based on set-up measurements after the irradiation. They used the knowledge of the distribution of random and systematic deviations. This procedure incorporated criteria for correcting the next patient set-up or not. A similar procedure was reported by Mitine et al. [11]. In these procedures, a decision was made based on one or more set-up measurements. These procedures aim at minimizing the number of deviations exceeding a specified value.

Another type of procedure was proposed by De Neve et al. [3] and Ezz et al. [5] where a correction is performed on-line during each daily irradiation. By using this procedure, one is able to correct both random and systematic deviations, but a major disadvantage related to such a procedure is the large increase in treatment time.

The purpose of the present study is to design and test a patient set-up verification procedure where decisions are made off-line after the irradiation. With this procedure, systematic deviations should be corrected as much as possible, thus resulting in a high overall set-up accuracy. This accuracy should be obtained with a minimum number of set-up measurements and corrections. In addition, the probability that a set-up has to be corrected repeatedly should be small for individual patients. It is also required that large systematic deviations of individual patients are detected and corrected at an early stage of the treatment.

2. Materials and methods

2.1. Set-up deviations

A set-up deviation is measured, in general, by comparing a portal image to a simulator image. In this study, however, we simulated set-up deviations with a computer to test our set-up verification procedure.

In the simulation, a set-up deviation \( d \), which can refer to a translation or a rotation, is modelled as the sum of a systematic deviation, \( \Delta \), and a random deviation \( \delta \) [2] (the Appendix gives an overview of the symbols). The random deviation during each irradiation is a sample of a normal distribution given by the function \( \phi(0, \sigma) \). The distribution of systematic deviations over all patients \( \phi(\Delta, M, \Sigma) \) is equal to the distribution of random deviations when there are no transfer errors; otherwise we have \( \Sigma > \sigma \) [2]. The parameter \( M \) specifies the mean value and \( \Sigma \) and \( \sigma \) the standard deviations of the corresponding distributions. The mean value of the distribution of random deviations is zero by definition. None of these parameters is supposed to be a function of time.

2.2 Measurement and correction of set-up deviations

Patient set-up deviations are measured during a consecutive number of fractions of the treatment. After each number of measurements \( N \), an average deviation, \( d_N \), is calculated.

If the average of the deviations is larger than an action level, the next set-up is corrected. The set-up is shifted with an intended magnitude equal to \( d_N \). Since the actual correction is not perfect in clinical practice, the accuracy of the correction is assumed to be normally distributed with a standard deviation \( \sigma_{\text{corr}} \) and an average value equal to zero.

2.3. Verification procedure and decision rules

The aim of the set-up verification procedure and the application of decision rules is to correct systematic deviations as much as possible. A set-up correction can be performed most accurately after a large number of measurements, since according to elementary statistics, \( d_N \) is a sample of a normal distribution with a standard deviation equal to \( \sigma/\sqrt{N} \) and a mean value equal to the systematic deviation. However, such a procedure is in contradiction with the need to correct large systematic set-up deviations as early as possible. A compromise between these two demands is reached by defining an action level which shrinks as the number of measurements increases. We used an action level equal to \( \alpha/\sqrt{N} \) [14] where the parameter \( \alpha \) is called the variable initial action level. The other parameter in the verification procedure, \( N_{\text{max}} \), is the maximum number of measurements used to calculate an average.

The procedure takes place in the first consecutive number of fractions (Figs. 1 and 2). After an irradiation, the set-up deviation is measured. This deviation is averaged with the preceding values of the deviations (if available). If the average deviation is smaller than the action level and the number of prescribed measurements \( N_{\text{max}} \) is not reached, the procedure is continued. On the other hand, if the average is larger than the action level, the patient set-up is corrected and the measurement...
Fig. 1. Flow diagram of the set-up verification procedure. After the measurement of the set-up deviation, the average deviation is calculated and tested against an action level. If the average is larger than the action level, the set-up is corrected, otherwise the procedure is continued. If the number of measurements since the last correction has reached a maximum value, the procedure is terminated. The symbols are explained in the text.

The verification procedure and the decision rule for set-up correction were tested by means of a computer simulation of the treatment of $10^4$ patients each receiving 30 fractions. Each set-up gets an initial systematic deviation, which is defined by a sample from the distribution of systematic deviations $\psi(\Delta,M,\Sigma)$; in another experiment $\Delta$ is taken to be constant. For each fraction a random deviation is added to the systematic deviation, which is a sample from the normal distribution of random deviations $\psi(0,0,\sigma)$.

2.5. Input parameters

To facilitate the application of the results for various clinical situations, the parameters $\alpha$, $\Sigma$ and $\sigma_{\text{corr}}$ were normalized with respect to $\sigma$. Several combinations of the parameters in the procedure ($\alpha$ and $N_{\max}$) were considered: $\alpha = 1\sigma$, $2\sigma$, $3\sigma$ and $4\sigma$ and $N_{\max} = 1, 2, 4$ and 8. A combination of these parameters will be denoted by $C(\alpha/\sigma,N_{\max})$.

Only values of $\Sigma$ equal to or larger than $\sigma$ are expected to occur in clinical practice [2]; we considered: $\Sigma = 1\sigma$, $2\sigma$ and $3\sigma$. Moreover, there is no reason to assume that $M$ is not equal to zero. In addition, the patient set-ups were simulated with a set of specific values of the systematic deviation $\Delta$ (1.5$\sigma$, 3.5$\sigma$ and 4.5$\sigma$), so not with a distribution.

The standard deviation of the accuracy of a set-up correction ($\sigma_{\text{corr}}$) was estimated by evaluating the actual clinical procedure for a correction. The set-up correction is performed by shifting the table or by drawing new skin marks. Each method yields an uncertainty, due to inaccurate table movements and the width of the skin marks, respectively. The error made in either case is assumed to have an uncertainty $\sigma_{\text{corr}}$ (one standard deviation) of the same magnitude as the standard deviation of the distribution of random deviations $\sigma$.

2.6. Output parameters

For various values of $\Sigma$, the resulting distribution of all the 30-$10^4$ set-up deviations $d$ (the resulting overall distribution) was obtained as a function of the input parameters $\alpha$ and $N_{\max}$. To simplify the interpretation of the results, the resulting overall accuracy was analyzed by considering one output parameter: the probability of occurrence of deviations larger than $3\sigma$, denoted by $P(d > 3\sigma)$.

The workload related to the procedure is determined by the weighted sum of the average number of measurements and corrections per patient. We do not specify the weights and investigated the two figures separately as a function of the parameters $\alpha$ and $N_{\max}$.

Since the average number of corrections per patient
3. Results

3.1. Resulting overall distribution

We investigated the effect of the verification procedure on the resulting distribution for various initial distributions of the systematic deviations (Fig. 3). As expected, for the three initial distributions of systematic deviations an increase of $N_{\text{max}}$ results in a better accuracy for nearly all values of $\alpha$.

This improvement becomes less for larger values of $N_{\text{max}}$. In the case of $\alpha = 1\sigma$, the increase of $N_{\text{max}} = 4$ to $N_{\text{max}} = 8$ results even in a slightly lower accuracy (Fig. 3), due to overcorrections.

The difference between the resulting accuracies with different initial distributions decreases as $N_{\text{max}}$ increases and $\alpha$ decreases. Thus, for appropriate values of $\alpha$ and $N_{\text{max}}$, the resulting overall distribution is nearly independent of the initial distribution of systematic deviations.

3.2. Workload

In order to study the workload in relation to a specific overall accuracy, we plotted the accuracy against the average number of measurements and corrections for various combinations $C(\alpha/\sigma, N_{\text{max}})$ (Fig. 4). The initial distribution of systematic deviation with $\Sigma = 2\sigma$ serves as an example. Without the application of the verification procedure, 18% of the deviations are larger than $3\sigma$. The best accuracy level that can be obtained is $P(d > 3\sigma) = 2.2\%$ (for $C(3,4)$).

In order to obtain a high accuracy with a low workload the appropriate parameter values should be determined, using Fig. 4. Consider, for instance, an overall accuracy level with set-up deviations larger than $3\sigma$ occurring in less than 5% of the set-ups, for short $P(d > 3\sigma) < 5\%$. Inspection of Fig. 4 reveals that generally this accuracy level, when obtained with a minimum number of measurements (Fig. 4a), can not be obtained with a minimum number of corrections (Fig. 4b) and vice versa. A trade-off between the number of measurements and corrections has to be made, as we will illustrate into more detail.

To get the specific accuracy level $P(d > 3\sigma) < 5\%$, the

![Fig. 3. A comparison between the effects of different initial distributions of the systematic deviation ($\Sigma = \sigma, 2\sigma$ and $3\sigma$) on the resulting accuracy as measured with the probability of deviations larger than $3\sigma$. The accuracy is shown as a function of $N_{\text{max}}$ for various values of $\alpha$.](image-url)
smallest number of measurements needed is obtained with the parameter combination C(1,1) (on average, 2.5 measurements per patient, Fig. 4a) and the corresponding number of corrections is, on average, 1.5 per patient (Fig. 4b). The required accuracy level is also obtained with C(2,2). In this case, the number of measurements increases to 3.2 on average and the number of corrections is now 0.9 on average, per patient. Furthermore, C(3,4) results in an acceptable overall accuracy but the average number of measurements and corrections per patient is 5.6 and 0.7, respectively. Inspection of Fig. 3 shows that for other initial distributions of the systematic deviations, the three parameter combinations (C(1,1), C(2,2) and C(3,4)) result also in the desired accuracy.

In conclusion, three parameter combinations result in an overall accuracy with deviations larger than 3σ in less than 5% of the cases, with either a minimum number of corrections (C(3,4)), a minimum number of measurements (C(1,1)) or a compromise between the number of corrections and measurements (C(2,2)). These three parameter combinations will be considered more extensively in the following subsections.

3.3. Repeated correction of patient set-ups

To study the probability of a large number of corrections per patient, the distribution of corrections per patient was calculated for the three optimal parameter combinations (C(1,1), C(2,2) and C(3,4). The standard deviation of the initial distribution of systematic deviations was Σ = 2σ (Fig. 5).

The larger the value of α, the smaller the probability that a patient has to be corrected repeatedly. For C(3,4) the probability that a patient has to be corrected more than once is 15.5%; for C(2,2) this is 22.5% and for C(1,1) this is 37.7%. Moreover, the spread of the distribution is much larger for C(1,1) than for C(3,4), as could be expected from the large number of corrections per patient for this parameter combination.

These results show that there is quite a large probability that a set-up has to be corrected repeatedly during a treatment for the parameter combination C(1,1). This unwanted situation is prevented somewhat with C(2,2), while with C(3,4) it is avoided as much as possible.
3.4. Large set-up deviations

Until now, only distributions of systematic deviations were considered. However, the procedure should also detect a large systematic deviation at an early stage of the treatment. Therefore, we studied the resulting distribution of the average deviation as a function of several initial systematic deviations (1.5α, 3.5α and 4.5α) for the three optimal combinations of parameters (C(1,1), C(2,2) and C(3,4); Fig. 6). The deviation in the first fraction was not included in the resulting average deviation.

With a moderate initial systematic deviation (1.5α) the procedure yields a resulting average deviation larger than 2α with a probability equal to 0.8% (C(1,1)) or smaller. With a large initial systematic deviation (3.5α or 4.5α) the procedure results in an average deviation larger than 2α with a probability of 2.2% (C(1,1)) or smaller. In general, the parameter combination with a small number of measurements (e.g., C(1,1)) yields a relatively large probability of a resulting large average deviation.

4. Discussion

A verification procedure with various initial action levels and a variable number of measurements and corrections has been investigated. The average number of measurements and corrections per patient can be minimized by varying the initial action level parameter α and the maximum number of measurements N_{max}. Three combinations of the parameters α and N_{max} are found which result in an accuracy level with deviations larger than 3α in less than 5% of the set-ups. With one of the parameter combinations, α = 2α and N_{max} = 2, the average number of measurements and corrections per patient is relatively small. The probability that a large initial systematic deviation is not corrected accurately, as well as the probability that a patient set-up has to be corrected often, is acceptably small.

However, the choice for a specific parameter combination, depends on the amount of work that is related to corrections and measurements. Expressed in terms of time or money, the workload related to a correction or measurement can be different from one institution to another. For instance, an institution with automatic portal image analysis procedures (e.g., [6]) could prefer performing measurements instead of corrections.

The choice of the parameters can also depend on the accuracy of the estimation of α. This estimation is crucial since the action level is proportional to α. To test the result for a different actual value of α, α_{act}, we calculated the number of measurements and corrections as well as the resulting accuracy (Table 1). As an example, consider the parameter combination α = 2α and

<table>
<thead>
<tr>
<th>α_{act}</th>
<th>No. measurements</th>
<th>No. corrections</th>
<th>P(d &gt; 3α_{act}) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>3.2</td>
<td>0.9</td>
<td>18.0</td>
</tr>
<tr>
<td>2α</td>
<td>5.5</td>
<td>2.9</td>
<td>18.0</td>
</tr>
<tr>
<td>α/2</td>
<td>2.3</td>
<td>0.2</td>
<td>18.0</td>
</tr>
</tbody>
</table>

The parameter combination is α = 2α and N_{max} = 2. The accuracy of the correction has a standard deviation equal to α_{act} and the distribution of systematic deviations has a standard deviation Σ equal to 2α_{act}.
\(N_{\text{max}} = 2\) and an initial distribution of systematic deviations with \(\Sigma = 2\sigma_{\text{act}}\). For simplicity we assumed that the accuracy of the correction is equal to \(\sigma_{\text{act}}\). From Table 1, it is clear that when the value of \(\sigma\) is underestimated, the number of corrections and measurements is much larger and the accuracy is hardly better than with a correct estimation of \(\sigma\). An overestimation of \(\sigma\) results in a lower accuracy and the number of measurements and corrections is substantially lower than with a correct estimation of \(\sigma\). When in a practical situation the value of \(\sigma\) is not known accurately, an unexpected increase in the workload is prevented by setting the value of \(\sigma\) (or \(\alpha\)) larger than was estimated. The consequence may be a decrease of the resulting overall accuracy.

The outcomes of the verification procedure for a certain parameter combination are also quite constant when some of the assumptions made in the present paper are not completely valid. We considered an overall mean unequal to zero for the parameter combination \(\alpha = 2\sigma\) and \(N_{\text{max}} = 2\) with \(\Sigma = 2\sigma\). Even with a large overall mean \(M = 2\sigma\), the verification procedure results in an accuracy with \(P(d > 3\sigma) = 4.0\%\) with a slight increase in the workload (3.5 measurements and 1.1 corrections per patient, on average). A more accurate correction than \(\sigma_{\text{corr}} = \sigma\) does not alter the outcomes dramatically, either. With a perfect correction (\(\sigma_{\text{corr}} = 0\)) the resulting accuracy is only slightly better (\(P(d > 3\sigma) = 2.9\%\)) and the corresponding workload somewhat lower (3.5 measurements and 1.1 corrections per patient, on average).

As a measure of the accuracy the probability of deviations larger than \(3\sigma\) (\(P(d > 3\sigma)\)) was used. Since the resulting overall distribution is shaped like a normal distribution, any other measure of the accuracy, e.g. \(P(d > 4\sigma)\), yields the same values for the optimal parameters.

In the present study it was assumed that systematic deviations do not vary during the course of the treatment. El-Gayed et al. [4] reported, however, the occurrence of significant time trends in 5 of the 20 patient set-ups that they considered. Hence, when time trends are expected, the verification procedure as described in the present paper should be modified. After measurements and subsequent corrections in the first consecutive fractions, regular checks should be done in order to determine and correct significant time trends.

Another assumption was that deviations can be described by the one-dimensional parameter \(d\). However, a measurement of the set-up deviation with portal imaging yields a two-dimensional vector. The one-dimensional model can be applied by considering the deviations separately in two orthogonal directions. A two-dimensional analysis, involving the length of the deviation vector, can be performed along the same lines as the analysis of this study. Preliminary results showed that the optimal values of the parameters are somewhat different. Therefore, further study is required to determine the optimal parameter combinations in this situation.

4.1. Other verification procedures

Based on the analyses of a large series of portal images, Mitine et al. [11] proposed a single check \((N_{\text{max}} = 1)\) with a corrective action when the deviation was larger than \(2\sigma\) (\(\alpha = 2\sigma\)). In this situation, the application of a single check verification procedure results in an accuracy which is quite strongly dependent on the initial distribution of systematic deviations (see Fig. 3). A single check verification procedure with other action levels result in either more corrections (\(\alpha = 1\sigma\), Fig. 4) or in a lower accuracy (\(\alpha = 3\sigma\), Fig. 3). Hence, a set-up verification procedure with only a single check should not be recommended [9].

Another verification procedure was proposed by Bijhold et al. [2]. They also used a decision rule with a shrinking action level. However, they assumed a known initial distribution of systematic deviations. We calculated the results of the application of the verification procedure of Bijhold et al. [2] for \(\Sigma = 2\sigma\). We introduced \(N_{\text{max}}\) in their procedure to be able to compare the results. The smallest value of \(N_{\text{max}}\) that satisfies the criterium \(P(d > 3\sigma) < 5\%\) is equal to 4, resulting in \(P(d > 3\sigma) = 2.9\%\). The average number of measurements and corrections per patient are 7.7 and 1.1, respectively. With the procedure of the current paper, the same accuracy can be obtained with a lower workload.

5. Conclusions

Presuming that the magnitude of the random deviations is known, the set-up verification procedure yields a high accuracy. The procedure uses an action level which is proportional to the estimated standard deviation of the distribution of random deviations and which shrinks as a function of the number of measurements. The resulting accuracy is nearly independent of the initial distribution of systematic deviations. By choosing appropriate values for the parameters of the procedure, the workload can be minimized. Moreover, using the procedure one can obtain a small probability that a set-up has to be corrected repeatedly while a large initial systematic deviation will be corrected at an early stage of the treatment.

Acknowledgements

We would like to thank Dr. Ali El-Gayed and Dr. Ben Mijnheer for useful suggestions during the preparation of this paper and Mario Pinkster for his assistance with software problems. This study was supported by the Dutch Cancer Society, NKB Project NKI 91-01.
References


Appendix

Notation

\[ d \], Set-up deviation equal to the sum of \( \delta \) and \( \Delta \);

\[ d_{av} \], Average of \( N \) values of \( d \);

\[ C(\alpha/\sigma, N_{\text{max}}) \], Specific combination of the parameters \( \alpha \) and \( N_{\text{max}} \);

\( N \), Number of measurements of the deviation since the first fraction or the last correction;

\( N_{\text{max}} \), Fixed number of measurements that should be performed after the first fraction or a correction;

\[ P(d > 3\sigma) \], Probability of deviations \( d \) larger than \( 3\sigma \);

\[ \psi(\delta, \mu, \sigma) \], Normal distribution, defined as:

\[
\frac{e^{-\left(\frac{(\delta - \mu)^2}{2\sigma^2}\right)}}{\sqrt{2\pi}\sigma}
\]

\( \alpha \), Parameter defining the action level: \( \alpha/\sqrt{N} \);  
\( \delta \), Random deviation;  
\( \Delta \), Systematic deviation;  
\( \sigma \), SD of the distribution of random deviations;  
\( \Sigma \), SD of the distribution of systematic deviations;  
\( M \), Mean of the distribution of systematic deviations.