The usefulness of an independent patient-specific treatment planning verification method using a benchmark plan in high-dose-rate intracavitary brachytherapy for carcinoma of the uterine cervix

Yutaka TAKAHASHI1,2,*, Masahiko KOIZUMI2, Iori SUMIDA3, Fumiaki ISOHASHI1, Toshiyuki OGATA1,2, Yuichi AKINO1, Yasuo YOSHIOKA1, Shintaro MARUOKA1, Shinichi INOUE4, Koji KONISHI1 and Kazuhiko OGAWA1

1Department of Radiation Oncology, Osaka University Graduate School of Medicine, 2-2 Yamada-oka, Suita, Osaka, Japan
2Division of Medical Physics, Oncology Center, Osaka University Hospital, 2-15 Yamada-oka, Suita, Osaka, Japan
3Department of Radiology, Osaka University Graduate School of Dental Medicine, 1-8 Yamada-oka, Suita, Osaka, Japan
4Department of Medical Technology, Osaka University Hospital, 2-15 Yamada-oka, Suita, Osaka, Japan
*Corresponding author. Tel: 81-6-6879-3482; Email: ytakahashi@radonc.med.osaka-u.ac.jp

(Received 2 March 2012; revised 15 May 2012; accepted 18 May 2012)

To develop an easy independent patient-specific quality assurance (QA) method using a benchmark plan for high-dose-rate intracavitary brachytherapy for cervix cancer, we conducted benchmark treatment planning with various sizes and combinations of tandem-ovoid and tandem-cylinder applications with ‘ideal’ geometry outside the patient. Two-dimensional-based treatment planning was conducted based on the Manchester method. We predicted the total dwell time of individual treatment plans from the air kerma strength, total dwell time and prescription dose of the benchmark plan. In addition, we recorded the height (dh), width (dw) and thickness (dt) covered with 100% isodose line. These parameters were compared with 169 and 29 clinical cases for tandem-ovoid or tandem-cylinder cases, respectively. With regard to tandem-ovoid cases, differences in total dwell time, dh, dt and dw between benchmark and individual plans were on average $-0.2\% \pm 3.8\%$, $-1.0 \text{ mm} \pm 2.6 \text{ mm}$, $0.8 \text{ mm} \pm 1.3 \text{ mm}$ and $-0.1 \text{ mm} \pm 1.5 \text{ mm}$, respectively. With regard to tandem-cylinder cases, differences in total dwell time, dhfr (the distance from tandem tip to tandem ring), dt and dw between benchmark and individual plans were on average $-1.5\% \pm 3.1\%$, $-1.5 \text{ mm} \pm 4.9 \text{ mm}$, $0.1 \text{ mm} \pm 1.0 \text{ mm}$ and $0.2 \text{ mm} \pm 0.8 \text{ mm}$, respectively. Of two cases, more than 13% differences in total dwell time were observed between benchmark plans and the clinical cases, which turned out to be due to the use of the wrong source position setting. These results suggest that our method is easy and useful for independent verification of patient-specific treatment planning QA.

Keywords: independent verification; treatment planning; Manchester method; benchmark plan; high-dose-rate intracavitary brachytherapy; uterine cervix

INTRODUCTION

Brachytherapy is an essential component of radiotherapy for the carcinoma of uterine cervix and is often combined with external beam radiation therapy (EBRT) for radical treatment. Several studies have suggested that control rates are significantly improved with EBRT and brachytherapy [1, 2]. High-dose-rate (HDR) remote afterloading intracavitary brachytherapy is widely used throughout Asia and Europe [3], and is becoming steadily more common in the USA [4].

The importance of independent verification of dosimetry prior to HDR brachytherapy treatment delivery has been recognized worldwide, and is specified in the guidelines of international regulatory agencies [5]. The Nuclear Regulation Commission (NRC) considers a 20% difference between the prescribed total dose and delivered dose to be a reportable medical event [5]. Thomadsen et al. identified 44 medical events in HDR brachytherapy between 1980 and 2001 in data from the NRC and International Atomic Energy Agency [6]. In fact, patients are often required to...
wait during treatment planning with an applicator inserted by a radiation oncologist, during which time errors and miscommunications can easily occur. This situation clearly indicates that patient-specific quality assurance (QA), including independent verification of treatment planning and confirmation of applicator geometry, should be done quickly and easily.

Many studies have reported independent verification methods for HDR brachytherapy treatment planning [7–16]. More recent reports have focused on the development of in-house software based on the AAPM TG 43 [17] formulation to calculate the dose at arbitrary points [13–15]. Although such software might be useful for the commissioning of treatment planning systems, human errors in individual treatment planning in clinical practice will not be identified due to the use of the same coordinate system, digitized applicator paths and dose point coordinates as those in the treatment planning system.

Although image-guided intracavitary brachytherapy has been enthusiastically investigated [18], treatment planning based on the Manchester method using two projection radiographs is still used [3]. One of the goals in intracavitary brachytherapy for carcinoma of the uterine cervix is to achieve the same level of consistency as the Manchester method. We have established the Osaka University Protocol based on the Manchester method with some modifications [19]. The goal of our institution is to achieve consistency with our protocol-based benchmark plans.

Here, we propose a very quick, simple and easy patient-specific independent verification method for Manchester method-based treatment planning using benchmark plans to detect human errors and evaluate the quality of the applicator geometry in the patient.

**MATERIALS AND METHODS**

**Creation of benchmark plans**

In this study, Fletcher-type (Fletcher-Williamson Asian-Pacific) tandem-ovoid and tandem-cylinder metal applicators (Nucletron International B.V., Veendaal, the Netherlands) were used. Various sizes and combinations of these applicators were constructed by one radiation oncologist with the ‘ideal’ applicator geometry outside the patient (Fig. 1a and b) and then reviewed by a medical physicist. We constructed eight kinds of tandem-ovoid and six kinds of benchmark plans (Table 1).

Figure 1c shows the ‘ideal’ geometry of a tandem-ovoid applicator used in our institution. Namely, a flange on the tandem tube is used at the origin, which is the cervical os

![Fig. 1. Creation of benchmark plans. (a) Scheme for the construction of tandem-ovoid; (b) tandem-cylinder applications; (c) typical dose distribution with the tandem-ovoid application.](https://academic.oup.com/jrr/article-abstract/53/6/936/970751)
and the tip of the ovoid is aligned with the origin (a flange on the tandem). Using these ideal applicator settings, treatment planning was performed with PLATO (Nucletron International B.V). Source dwell time was manually optimized based on the Manchester method as demonstrated by Tod and Meredith with minor modification [19–21]. Air kerma strength and total dwell time were then recorded. All benchmark plans were constructed by a medical physicist and reviewed by another medical physicist.

**Prediction of total dwell time for individual patient treatments**

Dose was calculated with the following formula introduced in AAPM TG43-U1 [22].

\[
D(r, \theta) = S_K \times \Lambda \times \frac{G_L(r, \theta)}{G_L(r_0, \theta_0)} \times g_L(r) \times F(r, \theta) \\
\times \text{dwell time}
\]

\[
= S_K \times A_b \times \text{dwell time} \times A
\]

where \(S_K\), \(\Lambda\), \(G_L\), \(g_L\) and \(F\) represent air kerma strength, dose rate constant, geometric function, radial dose function and anisotropy function, respectively. Here, we defined \(A\) as the product of \(\Lambda\), \(G_L\), \(g_L(r)\) and \(F(r, \theta)\).

The dose at the reference point in the benchmark plan is therefore calculated by the following:

\[
D(r, \theta)_b = A_b \times S_{K,b} \times \text{dwell time}_b
\]  

(1)

Similarly, the dose at point A in the individual plan is calculated by the following:

\[
D(r, \theta)_i = A_b \times S_{K,i} \times \text{dwell time}_i
\]  

(2)

If the ‘ideal’ tandem-ovoid geometry is achieved in the patient without any planning errors or misuse of the applicator, \(A_b\) is nearly equal to \(A_i\). Dwell time in the individual plan can therefore be predicted by the following formula from (1) and (2):

\[
\text{Dwell time}_i = \frac{D(r, \theta)_i \times S_{K,b} \times \text{dwell time}_b}{D(r, \theta)_b \times S_{K,i}}
\]  

(3)

where \(D(r, \theta)_b\) and \(D(r, \theta)_i\) represent the prescription dose of each treatment and the benchmark plan, respectively.

**Comparison of dose shape with that of the benchmark plans**

ICRU report 38 [23] recommends reporting the reference volume as well as total reference air kerma strength and absorbed dose at reference points. The reference volume is the volume encompassed by the reference isodose surface, which is represented by the major dimensions of the following:

(i) Height (dh), which is the maximum dimension along the tandem source measured on an ‘oblique’ sagittal plane;

(ii) Thickness (dt), which is the maximum dimension perpendicular to the tandem sources measured on a transverse plane;

(iii) Width (dw), which is the maximum dimension perpendicular to the tandem sources measured on a transverse plane.

In addition to the above parameters, we defined the dimensions of dh<sub>front</sub> and dh<sub>ext</sub>, which represent the distance from the 100% isodose line of the tip side of the tandem to the origin and that from the 100% isodose line of the connector side of the tandem to the origin, respectively (Fig. 2c).

Figure 2 shows the definitions of these parameters. For the tandem-cylinder, we recorded the additional parameters of dh<sub>front</sub> and dh<sub>ext</sub>, which represent the maximum dimension of the 100% isodose line of the tip side of the tandem to the tandem flange, and that of the connector side of the tandem to the tandem flange (Fig. 2c). These values were measured for individual treatment plans and compared with those of the benchmark plans.

**Analysis of clinical cases**

We retrospectively analyzed 168 and 29 clinical cases from 2009 through 2010 with a tandem-ovoid and tandem-cylinder, respectively. The difference in total dwell time between a benchmark plan and an individual treatment plan was calculated using the following formula:

\[
\text{Relative difference} (\%) = \frac{T_{\text{individual}} - T_{\text{benchmark}}}{T_{\text{benchmark}}} \times 100
\]

where \(T_{\text{benchmark}}\) and \(T_{\text{individual}}\) represent the total dwell time of the benchmark and individual plans, respectively. Differences in dose distribution shapes, including dh or dh<sub>front</sub>, dh<sub>ext</sub>, dt, and dw between the benchmark and
individual plans were calculated using the following formula:

\[
\text{Relative difference} = \frac{\text{shape}_{\text{benchmark}} - \text{shape}_{\text{individual}}}{\text{shape}_{\text{benchmark}}}
\]

Where \(\text{shape}_{\text{benchmark}}\) and \(\text{shape}_{\text{individual}}\) represent the dose distribution shapes of the benchmark and individual treatment plans, respectively.

Correlations between the differences in total dwell time and those in \(dh\), \(dt\) or \(dw\) among the benchmark and individual treatment plans were evaluated by Spearman’s rank correlation coefficient using Dr. SPSS II software (IMB, New York, USA).

Tolerance levels for total dwell time, \(dh\) or \(dh_{\text{front}}\), and \(dh_{\text{ext}}\), \(dt\), and \(dw\) were calculated by the following formula, which was first proposed by Venselaar et al. [24]

\[
\text{Tolerance level} = \text{mean deviation} \pm 1.96 \times SD
\]

**RESULTS**

**Tandem-ovoid cases**

*Prediction of total dwell time for individual treatment plans*

Figure 3 shows a histogram of differences in total dwell time between benchmark and individual treatment plans.

Differences averaged \(-0.2\% \pm 3.8\%\) (range, \(-13.3\%\)–\(-9.6\%\)), and exceeded 5% in 23 of 169 clinical cases.

**Comparison of the dose distribution shapes of individual treatment plans with those of the benchmark plans**

Figure 4 shows a histogram of differences in dose distribution shapes between the benchmark and individual treatment plans in tandem-ovoid applications. The differences in
dh, dt and dw between the benchmark and individual plans averaged \(-1.0\) mm \(\pm\) 2.6 mm (range, \(-8.6\) mm to +6.5 mm), 0.8 mm \(\pm\) 1.3 mm (range, \(-1.4\) mm to +5.2 mm) and \(-0.1\) mm \(\pm\) 1.5 mm (range, \(-5.1\) mm to +4.0 mm), respectively.

Regarding dh, 9 of 169 cases showed a difference between the benchmark and individual plans of greater than 5 mm. For dt and dw, in contrast, only one case showed a difference of more than 5 mm.

**Subset analysis of cases with large deviations between the benchmark and individual treatment plans**

We verified that all tandem-ovoid treatment plans were appropriately created without any planning errors, including wrong source position, wrong decay correction of source strength and inappropriate optimization or use of an unplanned size or combination of applicators. However, 24 of 169 cases had a >5% difference in total dwell time. To explain these differences, we investigated the correlation between differences in total dwell times and dh, dt and dw between the benchmark and individual plans. Figure 5a shows the relationship of differences in total dwell time (vertical axis) with those in dh (horizontal axis). Spearman’s rank correlation coefficient \((r_s = 0.836, P < 0.01)\) showed a strong relationship between the discrepancy in total dwell time and those in dh. In contrast, no correlations were found between the discrepancy in total dwell time and those in dt \((r_s = 0.371, P = 0.075)\) or dw \((r_s = 0.290, P = 0.149)\) (Fig. 5b and c).

**Tandem-cylinder cases**

Figure 6a shows differences in total dwell time between the benchmark and individual treatment plans. Differences averaged \(-1.5\% \pm 3.1\%\) (range, \(-13.0\%\) to +0.4\%), with 2 of 29 cases exceeding 11%.

Figure 6b shows the differences in dh\(_{\text{front}}\), dh\(_{\text{ext}}\), dt and dw between the benchmark and individual treatment plans. Differences averaged \(-1.5\) mm \(\pm\) 4.9 mm (range, \(-19.0\) mm to +4.0 mm), +1.8 mm \(\pm\) 5.2 mm (range, \(-2.2\) mm to +20.8 mm), +0.1 mm \(\pm\) 1.0 mm (range, \(-1.3\) mm to +4.3 mm) and +0.2 mm \(\pm\) 0.8 mm (range, \(-0.4\) mm to +3.9 mm), respectively. The differences in 2 of 29 cases, which also exceeded an 11% difference in total dwell time, exceeded \(-19\) mm and +20 mm for dh\(_{\text{front}}\) and dh\(_{\text{ext}}\), respectively. These cases were found to have been treated
with an unplanned tandem length, resulting in an incorrect setting for the source dwell positions in treatment planning.

Determination of tolerance limit
For tandem-ovoid cases, the tolerance level of total dwell time, dh, dt, and dw were −7.5% to +7.2, −6.0 mm to +4.1 mm, −1.8 mm to +3.4 mm and −3.0 mm to +2.8 mm, respectively (Fig. 7a).

For tandem-cylinder cases, two cases were excluded from the determination of tolerance limits because they were human error-related. Tolerance limits for total dwell time, dh_{front}, dh_{ext}, dt, and dw were −2.5% to +1.1%, −2.6 mm to +2.3 mm, −2.2 mm to +3.0 mm, −0.9 mm to +0.6 mm and −0.4 mm to +0.4 mm, respectively (Fig. 7b).

DISCUSSION
We used benchmark plans to establish a highly simple, easy, and fast independent treatment planning verification method for high-dose-rate intracavitary brachytherapy for carcinoma of the uterine cervix that requires no special skills such as developing TG43-based in-house software. Despite its great simplicity, analysis of a large number of clinical cases showed that our method was able to detect human error-related planning mistakes, and to evaluate the quality and consistency of applicator geometry.

The Manchester method, which was first suggested by Tod and Meredith in 1938, has been the most broadly used for the treatment of carcinomas of the uterine cervix, with some modifications from the original [20, 21]. They demonstrated an ‘ideal’ system in which loading patterns of milligrams of radium were determined based on the size of the tandem and ovoids to achieve as constant a dose rate at point A as possible, no matter what combination of applicators was used, and to ensure a suitable ratio between the intra-uterine and vaginal contribution [21]. This rule has been applied to high-dose-rate brachytherapy with an Ir-192 stepping source by modulating the weight of dwell times. In our institution, manual optimization in treatment planning is also based on the Manchester method with some modifications [19]. In addition, the applicator is set such that its geometry is consistent with the ‘ideal’ geometry specified in the benchmark plan. If there were any planning errors and inappropriate applicator setting, total
dwell time and dose distribution shapes, including dh, dt and dw, in individual treatment plans should agree with those of the benchmark plan. From these points, we established a method for the independent verification of patient-specific treatment planning QA by comparing benchmark plans with individual treatment plans.

Several other independent verification methods for individual treatment planning have been reported. Kumar et al. developed an in-house application that calculates the dose at arbitral points [13]. Lachaine et al. also developed an in-house application that achieves very fast calculation of point dose [14]. Such kind of applications are likely to be useful in the commissioning of treatment planning systems and partly also in individual treatment planning QA in terms of parameters such as source strength, treatment date and source table, which users input by themselves. However, because these applications use the same Cartesian coordinate system, digitized applicator paths and dose point coordinates as those in the treatment planning system, they are unable to detect human errors associated with the treatment planning process, such as setting of prescription point (Point A) with the wrong coordinate, the incorrect digitization of applicators, incorrect dose points or applicator points, improper magnification of simulation films, or use of an unplanned size or combination of applicators.

Several groups have previously proposed a method of checking total dwell time as a fraction of treatment length and dose prescription, or dose area index for planar implants and single catheter interstitial brachytherapy [8–11]. Recently, for example, Das et al. reported that total dwell time can be predicted from the reference volume covered with the prescribed dose (V100) in both single catheter and multiple catheter interstitial implants by the retrospective analysis of V100 from many clinical cases [12]. All these reports were focused on interstitial implants rather than intracavitary brachytherapy, however, and little information on intracavitary brachytherapy for carcinoma of the uterine cervix is available. In 1992, Thomadsen et al. demonstrated a method that assures the consistency of dwell times and dose distribution with previous treatment fractions and

---

**Fig. 6.** Differences in (a) total dwell time, and (b) dh\(_{\text{front}}\), dh\(_{\text{ext}}\), dt and dw between the benchmark plans and clinical cases in tandem-cylinder settings. White squares, dh\(_{\text{ext}}\); white diamonds, dh\(_{\text{front}}\); black triangles, dt; and crosses, dw.

**Fig. 7.** Tolerance levels of (a) tandem-ovoid and (b) tandem-cylinder.
previous patients [7]. Although our method is basically similar to their concepts, we created benchmark plans in which the ‘ideal’ geometry of the tandem-ovoid or tandem-cylinder can be achieved because these applications are constructed outside the patient’s body. Therefore the geometry of the applicator can be evaluated in every treatment by comparison with that of the ‘ideal’ geometry in the benchmark plan.

Although many independent verification methods have been reported, as described above, our present study is one of only a few to evaluate the usefulness of the method in a large number of clinical cases of intracavitary brachytherapy. In tandem-cylinder cases, two cases were found to have >11% differences in total dwell time between the benchmark and individual treatment plans, and >18 mm differences in dh\text{front} and dh\text{ext}. Review of these two cases showed that these differences were due to the unplanned use of tandem length, which resulted in the use of incorrect settings for the source dwell positions in treatment planning. The results clearly demonstrate that our method can easily identify such kinds of human error.

Regarding tandem-ovoid cases, a thorough review revealed no human-related planning errors. We next examined the reason why 24 cases of tandem-ovoid cases had >5% differences in total dwell time between the benchmark and individual plans. We found that these differences strongly correlated with differences in dh (Fig. 6), indicating that when the ovoid position shifts cranially compared with the benchmark plan, total dwell time decreases because the distance between the sources in the ovoids and point A becomes shorter. Conversely, if the ovoid position shifts caudally to the tandem flange compared with the benchmark plan, total dwell time increases, because the distance between sources in the ovoids and point A becomes larger. These facts indicate that our method is useful in not only finding human errors and software bugs but also in evaluating the quality of the applicator insertion technique. In other words, the evaluation of both the differences in total dwell time and dh could provide a good indicator for the quality of the applicator’s geometry.

We set tolerance limits for differences in total dwell time and dose shape between the benchmark and individual treatment plans. Ezell et al. reported that they set action levels for per-patient intensity modulated radiation therapy verification using confidence limits [25], as first proposed by Venselaar et al. [24]. If the confidence limit is established with sufficient points to provide good statistics, then the value of 1.96σ suggests that variations in excesses of the limit would occur about 5% of the time. We determined the tolerance limits by using confidence limits for total dwell time, dt, dh (dh\text{front} and dh\text{ext} for tandem-cylinder cases) and dw. To calculate tolerance limits, we excluded the two cases with >18% differences in tandem-cylinder cases to eliminate the effect of human error-related planning mistakes. These tolerance limits might be one indicator in the evaluation of individual treatment plans (i.e. rechecking of treatment planning, use of appropriate size of applicators, inappropriate applicator geometry, etc.).

One limitation of our study warrants mention, namely that our method is useful only for Manchester-based intracavitary brachytherapy. For carcinoma of the uterus, however, most treatment centers in the world have followed a traditional concept based on the Manchester loading patterns [3, 26]. Moreover, local control rates are significantly improved with EBRT and brachytherapy based on the Manchester method [1]. Our method therefore appears useful despite this limitation.

In conclusion, we established a highly simple, easy and quick independent verification method using benchmark plans for intracavitary brachytherapy based on the 2D-based Manchester method. Despite the great simplicity, our method can evaluate the quality of the applicator insertion technique, as well as identify human errors in treatment planning.

ACKNOWLEDGEMENTS

This study was supported by Grants-in-Aid for Scientific Research (22791194) from the Japan Society for the Promotion of Science. This study was also supported by the Japanese Society of Therapeutic Radiology and Oncology and Grants-in-Aid for Cancer Research (Nos. 23-3-3rd Term Cancer Control-General-007) from the Ministry of Health, Labor and Welfare of Japan.

REFERENCES

7. Thomadosen BR, Shahabi S, Stitt JA et al. High dose rate intracavitary brachytherapy for carcinoma of the cervix: The


