Radiotherapy and Oncology 135 (2019) 25–32

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original Article

Impact of inter- and intra-observer variabilities of catheter reconstruction on multi-catheter interstitial brachytherapy of breast cancer patients

Karoline Kallis, Theodoros Kaltsas, Stephan Kreppner, Michael Lotter, Vratislav Strnad, Rainer Fietkau, Christoph Bert *

Department of Radiation Oncology, Universitätsklinikum Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany

ARTICLE INFO

Article history:
Received 21 December 2018
Received in revised form 20 February 2019
Accepted 21 February 2019
Available online 5 March 2019

Keywords:
Breast cancer
Interstitial multi-catheter brachytherapy
Inter-observer variability

ABSTRACT

Purpose: The aim of this study was to evaluate inter- and intra-observer variabilities of catheter reconstruction and its dosimetric impact for multi-catheter interstitial breast cancer patients.
Methods and materials: In order to evaluate inter-observer variabilities (IOV) three medical physicists reconstructed the catheter traces of 13 patients. These manual reconstructions were further compared to the automatic reconstruction algorithm integrated into the planning system and one on purpose imprecise manual reconstruction. For intra-observer variabilities (IAV) repeated reconstructions of two physicists were compared for 13 patients. In total 426 catheters were considered. Keeping dwell times, dwell positions, the optimization and the normalization relative points constant the geometrical deviations between the corresponding dwell positions of the reference data set and the investigated reconstructions were evaluated. Also, the effect on the quality indices, such as coverage index (CI), dose non-uniformity ratio (DNR) or conformal index (COIN), and the exposure of the organs at risk were analyzed.
Results: Over all patients and all different catheter reconstructions considered for IOV a mean deviation between the corresponding dwell positions of 0.60 ± 0.35 mm was detected. The first observer had a mean deviation of 0.54 ± 0.32 mm, whereas the second observer yielded a mean deviation of 0.58 ± 0.37 mm. The length of the catheter traces varied in the mean by 0.51 ± 0.45 mm. The mean relative deviation of the CI, DNR, COIN, mean heart dose and mean lung dose varied by 0.27 ± 0.31%, 0.0027 ± 0.0025, 0.0036 ± 0.0033, 0.024 ± 0.019%, 0.05 ± 0.11%, respectively. The skin dose ($D_{0.2ccm}$) changed in the maximum 8.52%. On average IAV reached a deviation between the corresponding dwell positions of 0.49 ± 0.30 mm. IOVs and IAVs proved to be significantly different (Wilcoxon's test $p < 0.01$).
Conclusions: The study proved that a repeated reconstruction of the catheter traces does not lead to large geometrical deviations or to a significant change in the dose exposure. But the lack of ground truth makes the estimation of the quality of the reconstruction challenging. A precise reconstruction mapping the reality is a necessity for the planned dose delivery. With all considered reconstruction techniques reliable quality indices for the target and the organs at risk could be obtained.
 Radiotherapy and Oncology 135 (2019) 25–32

Multi-catheter interstitial brachytherapy (iBT) is a valid treatment technique of accelerated partial breast irradiation (APBI) [1,2]. In iBT usually 10–30 flexible plastic catheters are implanted into the breast to guide the radioactive Ir-192 source through the tissue. The dwell times (DTs) of defined dwell positions (DPs) can be adapted to get an optimized dose distribution with sufficient coverage of the planning target volume (PTV) while sparing dose to the surrounding tissue or organs at risk (OAR) and with a good cosmetic outcome [3–5]. In order to prevent gross dose deviations within the target volume and to ensure a safe, consistent and accurate treatment, quality assurance and minimization of uncertainties are necessary [6,7]. One important factor for the quality is an accurate reconstruction of the catheter traces and their precise representation in the treatment planning system [7].

Tsalpatouros et al. [8] showed an accuracy of around 1.00 mm for computed tomography (CT) based catheter reconstruction using the Baltas phantom and Hensley et al. [9] reported an inter-observer catheter reconstruction deviation in interstitial breast brachytherapy of 1.10 mm–2.80 mm between radiograph and CT based reconstruction using a phantom. The precision of

* Corresponding author at: Strahlenklinik, Universitätsklinikum Erlangen, Universitätsstraße 27, 91054 Erlangen, Germany.
E-mail address: christoph.bert@uk-erlangen.de (C. Bert).

https://doi.org/10.1016/j.radonc.2019.02.017
0167-8140/

the reconstruction is influenced by patient movement at the time of the CT, breathing artifacts, partial volumes effects, image distortion, the CT-resolution, and the observer [10]. The implanted catheters guide the radioactive source in time of the radiation and hence define the source path. An incorrect defined path leads to a wrong delivery of the treatment plan, as all the DPs and DTs are planned according to the previously defined catheter traces. The reconstruction of the catheters supposable influences the quality of the irradiation and serves as ground truth for additional quality assurance measures, such as the detection of errors prior to irradiation [6,11–13].

Various studies evaluated the influence of inter-observer variability (IOV) with respect to contouring especially for brachytherapy of the prostate [14–16]. For example Chicas-Sett et al. [15] found mean dose differences of $D_{0.1cc,}$ D_{1cc} and D_{2cc} of 10 Gy, 7.3 Gy and 6.6 Gy, respectively. Upreti et al. [17] investigated the IOV for target volume delineation on the dose quality indices for APBI using multi-catheter iBT. Considering different plans, they estimated a mean deviation of coverage index (CI) varying within the observers in the range of 0.756 ± 0.076% and 0.840 ± 0.070%. Further, the influence of the reconstructions was analyzed for gynecological application by various studies [18,19]. Tanderup et al. [19] analyzed the applicator reconstruction uncertainty for 20 cervical cancer patients by rotating and translating the applicator. The implants showed a great deviation for longitudinal displacement of 2.9% per millimeter. Hellebust et al. [18] investigated the reconstruction uncertainties of applicators for cervical cancer in a phantom study. They found a relative standard deviation in the dose at reference points for all reconstruction techniques of less than 2.8%.

The goal of the study was to determine the impact of inter-observer variability (IOV) and intra-observer variability (IAV) on the catheter reconstruction and its dosimetric consequence.

Materials and methods

Planning CT data set of twenty-six female breast cancer patients treated with multi-catheter iBT at the University Hospital Erlangen from November 2017 to June 2018 were selected for the evaluation. Selected details about the analyzed patient cohort are summarized in Table 1. HDR aPBI patients received 9 fractions of 3.8 Gy, HDR boost patients 2 fractions of 6 Gy, PDR aPBI 96 pulses of 0.52 Gy and PDR boost patients between 20 and 25 pulses with 0.7 to 0.8 Gy. In total 426 catheters were used for the evaluation. The direction of the catheters varied between cranial caudal and medial lateral. The catheter implantation is performed according to the GEC-ESTRO guidelines written by Strnad et al. [20].

Table 1
Patient specifications HDR = high-dose-rate; PDR = pulsed dose rate; PTV = planning target volume; aPBI = accelerated partial breast irradiation; DP = dwell position.

Parameter	Specification	Value
Number of patients	Total	26
	aPBI	11
	Boost	15
	Left breast	15
	Right breast	11
	HDR	19
	aPBI	10
	Boost	9
	PDR	7
	aPBI	1
	Boost	6
Age	Range	32–80 years
	Median	59 years
Recruitment Time		20/11/2017–25/06/2018
Number of Catheters	Range	13–22
	Mean	16
	Total	426
PTV	Range	26.71–116.86 ccm
	Mean	59.28 ccm
Number of active DP	Range	108–297
	Mean	193.70

Treatment protocol

After catheter implantation (flexible plastic implant tube, 6F, single leader, 30 cm, Elekta, Veenendaal, The Netherlands) with the help of a template, planning CT images are acquired and used for contouring and implant reconstruction. The CT images (axial thorax routine, 120 kV, B31s kernel, 50 cm field of view, SOMATOM Sensation open system, Siemens Healthcare GmbH, Erlangen, Germany) are captured in free breathing, without contrast agent and with a resolution of 0.38*0.38*2 mm^3. For all patients the planning target volume (PTV) and the organs at risk (OAR), such as the ipsilateral lung, heart, ribs and skin, are contoured and planned according to the GEC-ESTRO guidelines [21,22] and the instructions reported by Strnad et al. [7]. The treatment planning system Oncentra Brachy (Version 4.5.3, Elekta, Veenendaal, The Netherlands) is used for treatment planning and contouring.

In order to define the source path, the catheter traces are manually reconstructed by a medical physicist. Usually, first a rough catheter representation is delineated by defining corresponding catheter points, beginning at the connector end and continuing every third slice up to the distal end of the catheter. The suitable catheter region is characterized by the corresponding Hounsfield units ($-600 \leq HU \leq -200$). After a first iteration of reconstructing the catheters, the representation is fine-tuned and smoothed by considering all different orientations in the oblique slices (axial, sagittal and lateral). The aim is a continuous catheter reconstruction, such that no edges are visible and each catheter path is defined by a continuous smooth volume. The treatment planning system (TPS) interpolates linearly between the different defined points and allows bending of the catheters up to 30 degrees. After reconstruction of the catheter traces the DPs are defined and DTs are optimized such that all GEC-ESTRO requirements, see Table 2, are fulfilled and such that the PTV, extended by a margin of 4 mm, is homogeneously covered. The DP step size is set to 2.50 mm and usually the DPs are activated continuously.

Data analysis

For the estimation of possible inter-observer differences, the catheters of 13 patients out of 26 were reconstructed four times.

Table 2
GEC-ESTRO requirements $V_{implant}$ = implant volume; CI = coverage index; DNR = dose non-uniformity ratio [25]; PTV = planning target volume; COIN = conformal index [28]; MHD = mean heart dose; MLD = mean lung dose; V_{xGy} = relative volume receiving × Gy; D_{xccm} = relative dose given to most exposed × ccm of organ.

Region of Interest	Quality Requirements
Implant	$V_{implant} < 300\ cm^3$
	$DNR \leq 0.35$
PTV	$D_{90} \geq 100\%$
	$CI \geq 90\%$
	$V_{150} \leq 65$ ccm
	$V_{200} \leq 15$ ccm
	$COIN \geq 0.65$
Skin	$D_{0.2ccm} < 100\%$
	$D_{1ccm} < 90\%$
Ribs	$D_{0.1ccm} < 90\%$
	$D_{1ccm} < 80\%$
Heart	$MHD < 8\%$
	$D_{0.1ccm} < 50\%$
Lung ipsilateral	$MLD < 8\%$
	$D_{0.1ccm} < 60\%$

The clinical approved plan (CP) was compared to two reconstructions of medical physicists in training (MP1/MP2), one reconstruction without smoothing the catheter path (NP) and one reconstruction using the automatic reconstruction (AP) algorithm implemented in the TPS that is based on the implementation by Milckovic et al. [23]. The chosen parameter set was as follows: curvature radius equal to 1.00 cm; in-plane noise set to pixel spacing, points for extrapolation was chosen to 4; maximum relative spot size was set to 4.5; search radius per unit distance was defined as 1.0; minimum length was chosen to be 2.00 cm and the point spacing ranged from 0.10 cm to 3.00 cm. All AP and NP plans were established by MP2.

The evaluation of IAV is based on the remaining 13 patients. The division of the patient cohort was done according the availability, given that in 13 out of the 26 cases the clinical approved plan was established by MP1. MP1 and MP2 reconstructed the catheter paths twice with at least a difference of one week and the deviations between these plans were reported, in order to analyze IAV.

The DP distribution and the DTs were taken from the clinical plan and remained constant over all compared treatment plans. The Euclidean distances of the corresponding DPs and the change of the length of the catheters were analyzed. The length of the catheter is defined as the linear connection between the defined catheter points within the buttons, defining the start and end of the breast.

Further, the dosimetric quality indices, defined in the GEC-ESTRO guidelines [24], and the exposures of the OARs were considered, like conformal index (COIN), coverage index (CI) and dose non-uniformity ratio DNR. The DNR was defined as: "the proportion of the volume that is encompassed by a high clinically still tolerable dose at the volume that is covered by the reference isodose" [7,25]. A high change in the DNR would indicate hot spots within the dose distribution due to changes in the reconstruction. Further, the high-dose volumes ($V(D_{HD\_200})$, $V(D_{HD\_150})$) were defined as $V(D_{HD\_200}) = V(2.0{*}D_{ref})$ and $V(D_{HD\_150}) = V(1.5{*}D_{ref})$ [7]. Furthermore, not only the coverage and homogeneity of the target play an important role, but also the exposure of the OAR, such as skin, heart, ribs and lung. OAR Dxccm [%] defines the relative dose given to maximal exposed × ccm of the organ [24]. The mean lung dose (MLD) indicates the mean dose exposure of the defined ipsilateral lung contour, likewise the mean heart dose (MHD) is defined. All dose values of the OARs depend on the quality and completeness of the contour. On limiting factor of the completeness is the field of view of the planning CT. Therefore, the MHD and MLD should be considered with precaution, when comparing these values to different studies and among the patient cohort. As a result, the relative changes to the clinical plan are considered, using the same contour set.

All statistical computations were estimated with an in-house script implemented in R (R Foundation for Statistical Computing, version 3.3.2) [26]. The significance of the estimated deviations was proven by a grouped Wilcoxon–Mann–Whitney test with a significance level of 0.01. The maximal p-value of the investigated groups was used as total result. The deviations of each reconstruction were compared to each other. The Levene test with a significance level of 0.01 was used to prove the homogeneity of variances across groups.

Results

Inter-observer variability

All reconstructions were geometrically compared to the clinical approved treatment plan, thus all deviations state the difference to the CP and the corresponding DPs defined. In total 211 catheters and 2482 active DPs were considered. Fig. 1 shows an example of the different reconstructions and their influence on the dose distributions. In one case the automatic reconstruction could not be successfully performed, because of poor image quality. For all 13 patients and four catheter reconstructions a mean geometrical deviation of the corresponding DPs of 0.60 ± 0.35 mm was detected. Generally, MP1, MP2, AP, and NP had a median deviation with corresponding interquartile range of 0.48 ± 0.36 mm, 0.50 ± 0.39 mm, 0.62 ± 0.42 mm and 0.55 ± 0.37 mm, respectively. The geometrical variations yielded in the maximum 2.15 mm for MP1, 3.08 mm for MP2, 2.16 mm for NP and 2.95 mm for AP. All deviations ranged between 0.01 mm and 3.08 mm. The estimations were proven to be significantly different using the Wilcoxon test ($p < 0.01$). The length of the catheter traces varied in the mean by 0.51 ± 0.45 mm. AP resulted in the largest deviation with regards to length of 0.67 ± 0.59 mm. Over the considered parameters the MP1 yielded the smallest geometrical deviations to the CP. Nonetheless, MP2 was superior to NP and AP. Fig. 2a shows the distribution of the DPs' deviations of all the different users.

Further, the influence of different reconstruction schemes on the dosimetric quality indices was assessed. Fig. 3 gathers all the estimated information and all detailed results are summarized in Table 3. The coverage index (CI) of the patient cohort ranged from 90.02% to 97.88%. In the mean the CI varied by 0.27 ± 0.31% over all the estimations. MP1, MP2, NP, and AP had median deviation and interquartile range in the CI of 0.21 ± 0.25%, 0.25 ± 0.38%, 0.40 ± 0.28% and 0.17 ± 0.24%, respectively. MP1, MP2, NP, and AP reached a maximum CI deviation of 1.64%, 1.17%, 1.27% and 0.72%. The differences between NP, AP and MP1/2 proved to be not significant using the Wilcoxon test ($p > 0.01$). The dose non-uniformity ratio (DNR) ranged between 0.21 and 0.31 over the patient cohort and reached a mean value of 0.27 ± 0.03. On average, the DNR changed by 0.0034 ± 0.0025 when comparing the different reconstructions and different patients, see Fig. 3. MP1, MP2, NP, and AP had mean deviation in the DNR of 0.0040 ± 0.0033, 0.0034 ± 0.0027, 0.0039 ± 0.0035 and 0.0038 ± 0.0037, respectively. Over the selected patient cohort and reconstructions the mean COIN value was found to be 0.65 ± 0.08. Between the different reconstructions and patients the COIN varied between 0.001 and 0.011. On average the COIN for MP1, MP2, NP and AP varied not significantly (Wilcoxon's test, $p > 0.01$) by 0.0028 ± 0.0021, 0.0040 ± 0.0033, 0.0058 ± 0.0031 and 0.0026 ± 0.0025, respectively.

Intra-observer variability

In total 2554 DPs and 215 catheters of 13 patients were considered. In one out of these 13 cases, the heart contour was inaccurate, because of an insufficient field of view and hence could not be used for evaluation. Over all 13 patients a mean geometrical deviation of the corresponding DPs of 0.49 ± 0.30 mm was detected. The geometrical variations yielded in the maximum 2.54 mm for MP1 and 1.80 mm for MP2. All deviations ranged between 0.00 mm and 2.54 mm. The variance of the IOVs ($\sigma = 0.35$ mm) and the variance of the IAVs ($\sigma = 0.30$ mm) were proven to be significantly different using the Levene test ($p < 0.01$). Also the Wilcoxon test proved a significant difference of IOVs and IAVs ($p < 0.01$). The length of the catheter traces varied in the mean by 0.42 ± 0.37 mm. In comparison to MP2 (mean Euclidean distance = 0.57 ± 0.29 mm), MP1 (mean Euclidean distance = 0.42 ± 0.30 mm) yielded less intra-observer DPs' deviations. Fig. 2b shows the distribution of the user-dependent geometrical IAV. Fig. 2c compares IAVs to IOVs for catheter reconstructions estimated by MP2 using the IAV patient cohort. On average the DPs' deviations $MP2(1)_{MP2(2)}$, $MP2(1)_{MP1(1)}$, $MP2(2)_{MP1(1)}$, $MP2(1)_{MP1(2)}$ and $MP2(2)_{MP1(2)}$ were estimated to 0.57 ± 0.29 mm, 0.58 ± 0.34 mm, 0.54 ± 0.29 mm, 0.61 ± 0.37 mm and

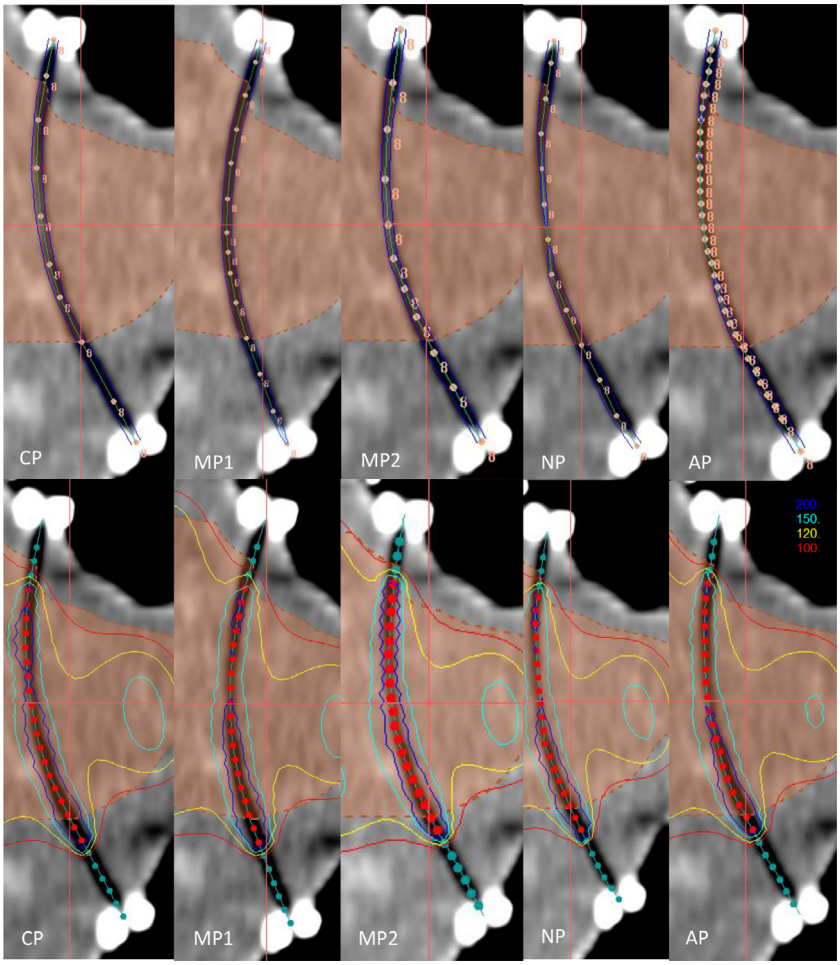

Fig. 1. Comparison of reconstruction techniques Clinical approved plan (CP), additional reconstructions by two medical experts in training (MP1, MP2), one reconstruction without smoothing the catheter traces (NP) and one reconstruction using the automatic reconstruction algorithm (AP) of patient P21 are presented. All images show the same image slice of the patient. In blue the projected catheter is visualized. The defined reconstruction points, defining the path of the catheter are presented in red in the first row. In orange the planning target volume is marked. The second row shows the dose distribution with the corresponding iso-dose lines (blue = 200%; turquoise = 150%; red = 100%; yellow = 120%). Also the active dwell positions (DPs) (red) and inactive DPs (cyan) are presented.

0.56 ± 0.30 mm, respectively. No fundamental difference in the behavior of IOV and IAV could be proven. All the repeated reconstructions yielded a mean error smaller than 1 mm, even considering all 26 patients. The geometrical IOVs average for each patient for MP2 ranged between 0.23 ± 0.12 mm for patient 25 and 0.84 ± 0.28 mm for patient 19, whereas the IAVs varied between 0.41 ± 0.37 mm for patient 26 and 0.64 ± 0.37 mm for patient 19.

The intra-observer influence on the dosimetric quality indices are summarized in Table 3 and the influence of the reconstruction on selected quality indices are shown in detail in Fig. 3. Here it is evident, that the variance of the dosimetric changes are not significantly different considering CI, DNR, COIN, Skin D_{1ccm} and MHD comparing IAV and IOV of MP1 (Levene's test, $p > 0.01$).

In summary, IOV and IAV resulted in geometrical deviation of 0.56 ± 0.34 mm on average considering all 26 patients. The influence of IOV and IAV on the selected dosimetric parameters was marginal and proved to be not significant.

Discussion

In iBT for breast cancer usually no additional positioning safety margins for the clinical target volume are added and steep dose gradients make the precise treatment planning and planned dose delivery essential [7]. The results showed that the catheter paths can be reproducible reconstructed by various users e.g. MP1, MP2, NP and AP (mean Euclidian distance = 0.60 ± 0.35 mm). For the estimation of IOV all reconstructions were compared to CP, which was considered to be the ground truth that is indeed also user dependent.

Comparing the deviations between MP1 (0.54 ± 0.32 mm) and observer MP2 (0.58 ± 0.37 mm) the mean reconstruction precision of MP2 is 0.04 mm worse considering IOV variations and the mean difference of 0.14 mm was estimated in the IAV within the second patient cohort. One weakness of the study is that the IOV results of NP and AP might be compromised by IAV variabilities, since all considered plans are created by the second observer.

Also the precise choice of the button center, defining the catheter tip, plays an important role, given that starting at this point the chain of dwell positions is continuously defined. If the first point has a rather large deviation all following dwell positions will be affected. For MP1, MP2, NP and AP the mean Euclidean distance to the first dwell position was estimated to be 0.49 ± 0.40 mm, 0.53 ± 0.39 mm, 0.52 ± 0.42 mm and 0.56 ± 0.30 mm, respectively.

We estimated an overall deviation between the corresponding DPs of 0.62 ± 0.36 mm using the automatic reconstruction algorithm. Similar results (mean geometrical errors = 0.87 ± 0.35 mm) were proven by Milickovic et al. [23] using 30 different clinical

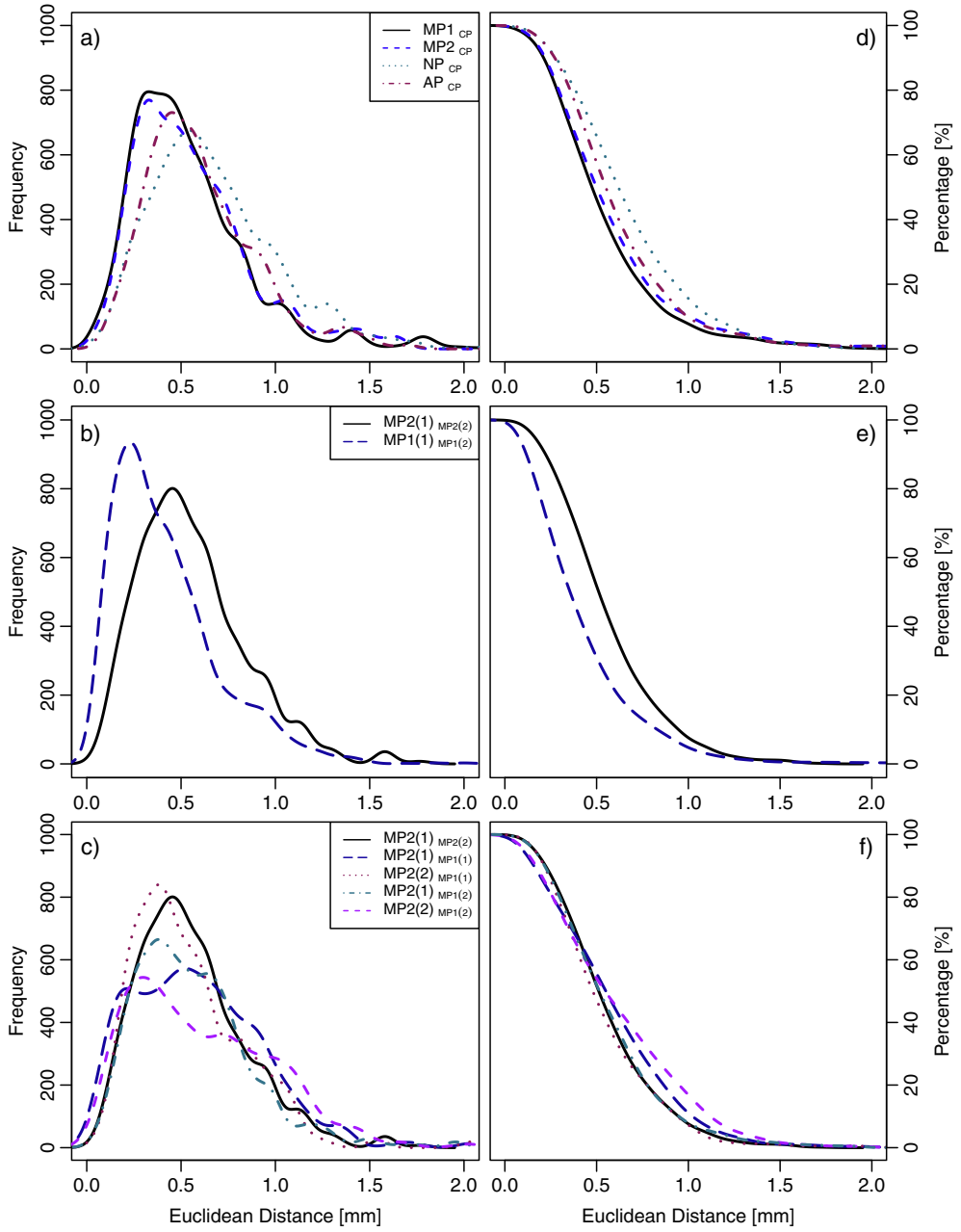

Fig. 2. Geometrical deviations For each reconstruction the deviations over all considered patients to the reference plan were plotted ($reconstruction_{reference}$). On the left side the density distribution of the Euclidean distance between the corresponding DPs is shown. A Gaussian kernel was used for smoothing. Subfigure a) presents the IOV. Subfigure b) shows the IAV for MP1 and MP2. Subfigure c) compares inter- and intra-observer deviations for all MP2 estimations. Subfigure d) shows the (1-cumulative deviations) distribution of the calculated distributions for IOV. Subfigure e) presents the (1-cumulative deviations) distribution of the calculated distributions for IAV comparing MP1 and MP2. Subfigure f) presents the (1-cumulative deviations) distribution of the comparison of inter and intra-observer deviations for all MP2.

implants. They also reported a time advantage of using the automatic reconstruction in 90% of the considered cases. They estimated a duration of 21.4 s per catheter for the automatic reconstruction in comparison to 684 s per catheter for using the manual reconstruction. In comparison we evaluated an average duration for the manual reconstruction of 139.19 ± 47.54 s per catheter considering 94 catheter traces. However, in our evaluation user intervention was still necessary using the automatic reconstruction because of blurred CT images, additional air holes within the tissue and in order to define the start/end point within the button center. Thus, our experience showed that the time advantage of using the automatic reconstruction algorithm depends on the image quality and could be more time consuming than the manual reconstruction in some cases.

The considered catheter length is calculated as linear combination between the chosen catheter points and hence the precision is influenced by the amount of defined points. That means if fewer points define curved catheters, the calculated length of the catheter might be smaller than the representation of the same catheter with an increased number of points. The large increase in catheter length seen in AP of 0.67 ± 0.59 mm could be impaired by the large amount of catheter points. This difference can be seen in Fig. 1. Even with a relatively large geometrical deviation the automatic reconstruction has a low impact on the dosimetric parameters, like

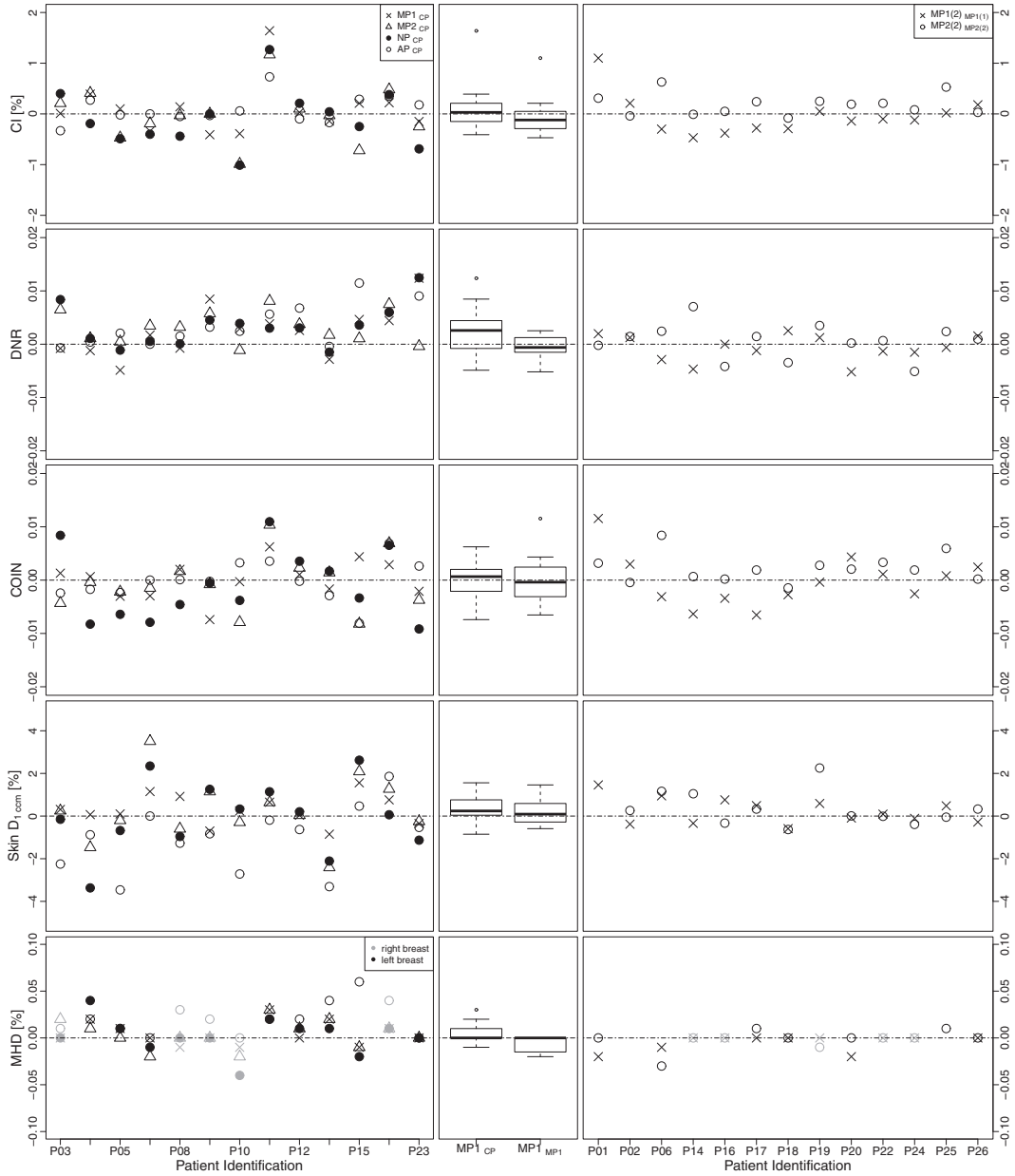

Fig. 3. Quality indices For selected quality indices, the influence of the different reconstructions is shown. Each value is normalized to the reference plan; hence the image shows relative deviations. MP1 and MP2 represent medical physics experts in training, NP the not smoothed reconstruction and AP the automatic reconstruction using the TPS. For each patient the inter- and intra-observer deviations are represented. Inter-observer variations are shown on the left side. On the right side intra-observer differences are presented and in the middle the IOV of MP1 are compared to IAV.

the CI, DNR and COIN, but the exposure of the OAR increased. A small shift of the catheter trace could be beneficial for the coverage but overall the estimation of the user driven clinical plan tries to optimize more parameters than the CI, hence even if an increase can be estimated it does not allow the conclusion that the treatment plan might be positively influenced in the clinical situation. Therefore, it is important to consider all dosimetric parameters to evaluate the influence of geometrical changes.

Non-smoothed catheter traces (NP) show the effect of an imprecise reconstruction and thus lead to false relative DPs. Hence, optimizing the plan with an imprecise reconstruction might fulfill all dosimetric requirements but does not account for the true situation since the DPs defined during treatment planning differ from the actual positions during treatment. Furthermore, an imprecise reconstruction also had an effect on the defined catheter length (0.56 ± 0.39 mm). Since all DPs are explicitly defined through the absolute DPs' distance to the connector, a false definition of the catheter length would lead to shifted DPs in the TPS. The defined length of the catheter influences the precision of the irradiation and proved to be one crucial parameter for the planned dose delivery [27]. Therefore, it seemed to be essential to define the correct start and end point of the catheters and to create a rather smooth catheter representation, although the reconstructions did not show a significant impact on the dosimetric quality indices.

The lack of a ground truth made it difficult to evaluate the quality of the reconstruction. In our study we used the CP as comparison, since this plan was verified and approved by an experienced physician and an experienced medical physicist.

Table 3
Comparison of geometrical and dosimetric deviations considering all estimated reconstructions The table summarizes all the estimations considering all plans and all patients. For each reconstruction the mean deviation to the reference plan and the standard deviation is presented. The significance of the group difference was estimated with the Wilcoxon–Mann–Whitney test with a significance level of 0.01. $\Delta MP1_{CP}$ = deviation of observer 1 to clinical plan; $\Delta MP2_{CP}$ = deviation of observer 2 to clinical plan; ΔNP_{CP} = deviation of not smoothed reconstruction to clinical plan; ΔAP_{CP} = deviation of automatic reconstruction to clinical plan; $\Delta MP1(1)_{MP1(2)}$ = IAV of observer 1; $\Delta MP2(1)_{MP2(2)}$ = IAV of observer 2; CI = coverage index; DNR = dose non-uniformity ration [25]; COIN = conformity index [28]; MHD = mean heart dose; MLD = mean lung dose.

Quality Index	$\Delta MP1_{CP}$	$\Delta MP2_{CP}$	ΔNP_{CP}	ΔAP_{CP}	$\Delta MP1(1)_{MP1(2)}$	$\Delta MP2(1)_{MP2(2)}$	Significance
Distance [mm]	0.54 ± 0.32	0.58 ± 0.37	0.67 ± 0.33	0.62 ± 0.36	0.42 ± 0.30	0.56 ± 0.29	$p < 0.01$
Length [mm]	0.51 ± 0.40	0.42 ± 0.37	0.56 ± 0.39	0.67 ± 0.59	0.30 ± 0.30	0.53 ± 0.39	$p < 0.01$
CI [%]	0.32 ± 0.42	0.39 ± 0.37	0.44 ± 0.36	0.20 ± 0.20	0.28 ± 0.28	0.20 ± 0.19	$p > 0.01$
DNR	0.0040 ± 0.0033	0.0034 ± 0.0027	0.0038 ± 0.0035	0.0038 ± 0.0037	0.0020 ± 0.0015	0.0025 ± 0.0021	$p > 0.01$
COIN	0.0028 ± 0.0021	0.0040 ± 0.0033	0.0058 ± 0.0031	0.0026 ± 0.0025	0.0027 ± 0.0030	0.0025 ± 0.0024	$p > 0.01$
PTV D90 [%]	0.43 ± 0.62	0.52 ± 0.43	0.57 ± 0.47	0.36 ± 0.37	0.37 ± 0.37	0.34 ± 0.31	$p > 0.01$
Skin $D_{0.2ccm}$ [%]	1.60 ± 1.68	2.35 ± 2.48	2.42 ± 1.79	1.31 ± 1.79	1.08 ± 1.03	0.90 ± 0.65	$p > 0.01$
Skin $D_{1.0ccm}$ [%]	0.59 ± 0.46	1.09 ± 1.05	1.26 ± 1.05	1.42 ± 1.18	0.51 ± 0.39	1.41 ± 3.10	$p > 0.01$
Heart MHD[%]	0.0092 ± 0.0095	0.0115 ± 0.0099	0.0131 ± 0.0138	0.0208 ± 0.0180	0.0512 ± 0.1721	0.0050 ± 0.0.009	$p > 0.01$
Heart $D_{0.1ccm}$ [%]	0.07 ± 0.08	0.10 ± 0.10	0.11 ± 0.12	0.19 ± 0.21	0.09 ± 0.15	0.07 ± 0.11	$p > 0.01$
Lung MLD [%]	0.0085 ± 0.0069	0.0138 ± 0.015	0.1846 ± 0.4656	0.0331 ± 0.034	0.0108 ± 0.0119	0.054 ± 0.0066	$p > 0.01$
Lung $D_{0.1ccm}$ [%]	0.62 ± 0.96	0.32 ± 0.34	0.42 ± 0.47	0.66 ± 0.66	0.29 ± 0.37	0.20 ± 0.17	$p > 0.01$
Ribs $D_{0.1ccm}$ [%]	0.63 ± 0.55	0.41 ± 0.42	0.56 ± 0.57	0.94 ± 0.69	0.89 ± 1.01	0.37 ± 0.50	$p > 0.01$
Ribs $D_{1.0ccm}$ [%]	0.30 ± 0.26	0.31 ± 0.31	0.46 ± 0.49	0.67 ± 0.44	0.67 ± 1.08	0.19 ± 0.15	$p > 0.01$

Treatment plan optimization depended on the representation of the catheter paths and hence the catheter reconstruction influenced the planned dose delivery. This means that even with an imprecise reconstruction of the catheter traces the dosimetric quality indices could yield sufficient results, but the treatment delivery would be corrupted. We also tried to minimize the risk of systematic errors by choosing a slice thickness of 2 mm, by verifying the reconstructions and by following in-house guidelines. Likewise, for cervical cancer Tanderup et al. [19] stated that all random errors can be minimized by a small slice thickness of the CT ($\leq$5 mm) and having clear guidelines. They estimated deviations of the quality indices of 5–10% for reconstruction of cervical applicators.

The differences between IOV and IAV variabilities seem to be marginal (<0.1 mm) although the variances of the groups were proven to be significant different (Levene's test $p < 0.01$). Considering, the two independent patients cohort the mean value of IAV (mean Euclidian distance = 0.49 ± 0.30 mm), was smaller than the IOV (mean Euclidian distance = 0.56 ± 0.35 mm). Comparing the IAV and IOV variations within one group no such effect could be proven, which leads us to the conclusion that the patient choice has an influence on the variances rather than the difference between IAV and IOV.

The correctness of the estimated dose exposure on the OAR strongly depends on the quality of the contours. Imperfections lead to possible false dose peaks which lead to inadequate values considering small volumes. Also, the value of the MHD or the MLD is often corrupted due to a too small CT field of view. Especially, the heart contour for breast cancer patients, treated on the right side, is often incomplete in planning CT images. Thus, these values have to be considered with regard to the 3D contour with focus on relative comparisons among the 4 observers.

The influence on the dose received by the OAR did not change significantly within the different reconstructions. Even NP, representing the worst case, did not yield major deviations. The CI, DNR and COIN as quality indices show the worth of the calculated plan and the study lead to the impression that these values are rather influenced by the optimization of the DTs than the reconstructions. In contrast the IOV in target definition showed a significant impact on the dose quality indices (Wilcoxon's test $p < 0.05$) [17]. Upreti et al. [17] reported maximum percentage variations of e.g. the CI considering five observers of 1.2%, 5.2%, 6.1%, 4.9% and 6.3%. Overall, in our study the CI changed for MP1, MP2, NP and AP in the maximum 1.64%, 1.17%, 1.27% and 0.93%, respectively. In summary, no large deviations within the dosimetric quality indices were found and no significance proven (Wilcoxon's test $p > 0.05$).

With the current settings: CT resolution (0.53*0.53*2.00 mm), 6F catheters, margin (4 mm), number of catheters (mean number of catheters = 16), DP step-size (2.5 mm), optimization algorithm (TG-43) and inter-catheter spacing (template 12–20) the treatments plans seem to be robust to small geometrical variations considering the dosimetric quality indices. Further, all our results have to be considered regarding these settings which make them in detail hard to compare to other studied e.g. reconstruction uncertainties of gynecological applicators [18,19].

The results show that catheter paths can be reproducible and reliably reconstructed. The different reconstructed catheters showed negligible inter- and intra-observer deviations, which never would lead to a under dosage of the target volume. Also the exposure on the OAR was not corrupted by observer variations. No major differences between inter-observer and intra-observer variations could be proved.

Acknowledgements

The present work was performed in fulfillment of the requirements for obtaining the degree Dr. rer. biol. hum. of the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU).

Conflict of interest

None. The department has a research framework agreement with Elekta.

References

[1] Strnad V et al. 5-Year results of accelerated partial breast irradiation using sole interstitial multicatheter brachytherapy versus whole-breast irradiation with boost after breast-conserving surgery for low-risk invasive and in-situ carcinoma of the female breast: a randomised, phase 3, non-inferiority trial. Lancet 2016;387:229–38.
[2] Krug D. Accelerated partial breast irradiation using interstitial multicatheter brachytherapy: a valid treatment option for breast cancer patients with a low-risk profile? Strahlenther Onkol 2016;192:127–9.
[3] Njeh CF, Saunders MW, Langton CM. Accelerated Partial Breast Irradiation (APBI): a review of available techniques. Radiat Oncol 2010;5:1–28.
[4] Lettmaier S et al. Radiation exposure of the heart, lung and skin by radiation therapy for breast cancer: a dosimetric comparison between partial breast irradiation using multicatheter brachytherapy and whole breast teletherapy. Radiother Oncol 2011;100:189–94.
[5] Soror T et al. Cosmetic changes following surgery and accelerated partial breast irradiation using HDR interstitial brachytherapy: evaluation by a

multidisciplinary/multigender committee. Strahlenther Onkol 2017;193:367–74.
[6] Bert C, Kellermeier M, Tanderup K. Electromagnetic tracking for treatment verification in interstitial brachytherapy. J Contemp Brachyther 2016;8:448–53.
[7] Strnad V, Pötter R, Kovács G. Practical handbook of brachytherapy. Bremen: UNI-MED Verlag; 2014.
[8] Tsalpatouros A et al. CT-based software for 3-D localization and reconstruction in stepping source brachytherapy. IEEE Trans Inf Technol Biomed Eng 1997;1: 229–42.
[9] Hensley F et al. Precision and dose conformation in the CT-based reconstruction of interstitial brachytherapy. Med Phys 2003;30:1431.
[10] Kirisits C et al. Review of clinical brachytherapy uncertainties: analysis guidelines of GEC-ESTRO and the AAPM. Radiother Oncol 2014;110:199–212.
[11] Kellermeier M et al. Electromagnetic tracking (EMT) technology for improved treatment quality assurance in interstitial brachytherapy. J Appl Clin Med Phys 2017;18:211–22.
[12] Kallis K et al. Introduction of a hybrid treatment delivery system used for quality assurance in multi-catheter interstitial brachytherapy. Phys Med Biol 2018;63:095008.
[13] Kallis K et al. Quality assurance by electromagnetic tracking in interstitial multi-catheter HDR brachytherapy using a hybrid afterloading system. Strahlenther Onkol 2018;194:S20–1.
[14] Crook J et al. Interobserver variation in postimplant computed tomography contouring affects quality assessment of prostate brachytherapy. Brachytherapy 2002;1:66–73.
[15] Chicas-Sett R et al. Interobserver variability in rectum contouring in high-dose-rate brachytherapy for prostate cancer: a multi-institutional prospective analysis. Brachytherapy 2018;17:208–13.
[16] De Brabandere M et al. CT- and MRI-based seed localization in postimplant evaluation after prostate brachytherapy. Brachytherapy 2013;12:580–8.
[17] Upreti RR et al. Impact of inter-observer variations in target volume delineation on dose volume indices for accelerated partial breast irradiation with multi-catheter interstitial brachytherapy. Radiother Oncol 2018;129: 173–9.
[18] Hellebust TP et al. Reconstruction of a ring applicator using CT imaging: impact of the reconstruction method and applicator orientation. Phys Med Biol 2007;52:4893–904.
[19] Tanderup K et al. Consequences of random and systematic reconstruction uncertainties in 3D image based brachytherapy in cervical cancer. Radiother Oncol 2008;89:156–63.
[20] Strnad V et al. ESTRO-ACROP guideline: interstitial multi-catheter breast brachytherapy as Accelerated Partial Breast Irradiation alone or as boost – GEC-ESTRO Breast Cancer Working Group practical recommendations. Radiother Oncol 2018;128:411–20.
[21] Strnad V et al. Recommendations from GEC ESTRO Breast Cancer Working Group (I): target definition and target delineation for accelerated or boost Partial Breast Irradiation using multicatheter interstitial brachytherapy after breast conserving closed cavity surgery. Radiother Oncol 2015;115:342–8.
[22] Major T et al. Recommendations from GEC ESTRO Breast Cancer Working Group (II): target definition and target delineation for accelerated or boost partial breast irradiation using multicatheter interstitial brachytherapy after breast conserving open cavity surgery. Radiother Oncol 2016;118:199–204.
[23] Milickovic N et al. Catheter autoreconstruction in computed tomography based brachytherapy treatment planning. Med Phys 2000;27:1047–57.
[24] Strnad V et al. ESTRO-ACROP guideline: interstitial multi-catheter breast brachytherapy as Accelerated Partial Breast Irradiation alone or as boost – GEC-ESTRO Breast Cancer Working Group practical recommendations. Radiother Oncol 2018;128:411–20.
[25] Saw CB, Suntharalingam N. Quantitative assessment of interstitial implants. Int J Radiat Oncol Biol Phys 1991;20:135–9.
[26] R Developement Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing; 2008.
[27] Thomadsen B, Das R. Quality assurance for breast brachytherapy. In: Wazer D, Arthur D, Vicini F, editors. Accelerated partial breast irradiation: techniques and clinical implementation. Berlin Heidelberg: Springer-Verlag; 2006. p. 179–245.
[28] Baltas D et al. A conformal index (COIN) to evaluate implant quality and dose specification in brachytherapy. Int J Radiat Oncol Biol Phys 1998;40:515–24.