Published in final edited form as:

Int J Radiat Oncol Biol Phys. 2014 July 1; 89(3): 674–681. doi:10.1016/j.ijrobp.2014.03.005.

Dosimetric Consequences of Interobserver Variability in Delineating the Organs at Risk in Gynecologic Interstitial Brachytherapy

Antonio L. Damato, PhD^{*}, Kanopkis Townamchai, MD^{*}, Michele Albert, MD[†], Ryan J. Bair, MD[‡], Robert A. Cormack, PhD^{*}, Joanne Jang, MD, PhD[§], Arpad Kovacs, MD, PhD[‡], Larissa J. Lee, MD^{*}, Kimberley S. Mak, MD^{||}, Kristina L. Mirabeau-Beale, MD^{||}, Kent W. Mouw, MD, PhD^{||}, John G. Phillips, MD^{||}, Jennifer L. Pretz, MD^{||}, Andrea L. Russo, MD^{||}, John H. Lewis, PhD^{*}, and Akila N. Viswanathan, MD, MPH^{*}

^{*}Department of Radiation Oncology, Dana-Farber Cancer Institute/Brigham and Women's Hospital, Boston, Massachusetts

[†]Department of Radiation Oncology, Saint Anne's Hospital Regional Cancer Center, Fall River, Massachusetts

[‡]Department of Radiology, Brigham and Women's Hospital, Boston, Massachusetts

[§]Department of Radiation Oncology, Beth Israel Deaconess Medical Center, Boston, Massachusetts

^{||}Harvard Radiation Oncology Program, Harvard Medical School, Boston, Massachusetts

Abstract

Purpose—To investigate the dosimetric variability associated with interobserver organ-at-risk delineation differences on computed tomography in patients undergoing gynecologic interstitial brachytherapy.

Methods and Materials—The rectum, bladder and sigmoid of 14 patients treated with gynecologic interstitial brachytherapy were retrospectively contoured by 13 physicians. Geometric variability was calculated using κ statistics, conformity index (CI_{gen}), and coefficient of variation (CV) of volumes contoured across physicians. Dosimetric variability of the single-fraction $D_{0.1cc}$ and D_{2cc} was assessed through CV across physicians, and the standard deviation of the total EQD2 (equivalent dose in 2 Gy per fraction) brachytherapy dose (SD^{TOT}) was calculated.

Results—The population mean $\pm$ 1 standard deviation of κ , CI_{gen} and volume CV were, respectively: 0.77 ± 0.06 , 0.70 ± 0.08 and $20\% \pm 6\%$ for bladder; 0.74 ± 0.06 , 0.67 ± 0.08 and $20\% \pm 5\%$ for rectum, and 0.33 ± 0.20 , 0.26 ± 0.17 and $82\% \pm 42\%$ for sigmoid. Dosimetric variability

© 2014 Elsevier Inc. All rights reserved.

Reprint requests to: Antonio L. Damato, PhD, Brigham and Women's Hospital, Department of Radiation Oncology, 75 Francis St, ASB1, L2, Boston, MA 02115. Tel: (617) 525-7242; adamato@lroc.harvard.edu.

K.T.'s present address is: Bhumibol Adulyadej Hospital, Bangkok, Thailand.

A.K.'s present address is: Department of Radiography, University of Pecs, Kaposvar, Hungary.

Conflict of interest: none.

Supplementary material for this article can be found at www.redjournal.org.

was: for bladder, $CV = 31\% \pm 19\%$ ($SD^{TOT} = 72 \pm 64$ Gy) for $D_{0.1cc}$ and $CV = 16\% \pm 10\%$ ($SD^{TOT} = 9 \pm 6$ Gy) for D_{2cc} ; for rectum, $CV = 11\% \pm 5\%$ ($SD^{TOT} = 16 \pm 17$ Gy) for $D_{0.1cc}$ and $CV = 7\% \pm 2\%$ ($SD^{TOT} = 4 \pm 3$ Gy) for D_{2cc} ; for sigmoid, $CV = 39\% \pm 28\%$ ($SD^{TOT} = 12 \pm 18$ Gy) for $D_{0.1cc}$ and $CV = 34\% \pm 19\%$ ($SD^{TOT} = 4 \pm 4$ Gy) for D_{2cc} .

Conclusions—Delineation of bladder and rectum by 13 physicians demonstrated substantial geometric agreement and resulted in good dosimetric agreement for all dose-volume histogram parameters except bladder $D_{0.1cc}$. Small delineation differences in high-dose regions by the posterior bladder wall may explain these results. The delineation of sigmoid showed fair geometric agreement. The higher dosimetric variability for sigmoid compared with rectum and bladder did not correlate with higher variability in the total brachytherapy dose but rather may be due to the sigmoid being positioned in low-dose regions in the cases analyzed in this study.

Introduction

Interstitial brachytherapy for the treatment of some gynecologic malignancies was first reported in 1913 (1). Historically, planning was performed according to source distribution rules independent of the surrounding anatomy (2, 3). Absorbed dose at reference points was used to estimate the dose to the organs at risk (OARs) (4). In modern times, 3-dimensional (3D) image-guided interstitial brachytherapy with computed tomography (CT) (5, 6) or magnetic resonance imaging (MRI) (7, 8) has been established (9); these tools facilitate dose escalation to the tumor while limiting the dose to the OARs. Moreover, 3D image guidance has been associated with increased local control and reduced morbidity in gynecologic brachytherapy (10).

With the advent of 3D planning, OAR dose-volume histogram parameters such as $D_{0.1cc}$ and D_{2cc} (the minimum dose in the most irradiated $0.1\text{cm}^3/2\text{cm}^3$ volume of the organ) (11) have been shown to be predictive of toxicity (11-14). All dose-volume histogram calculations rely on the assumption that the physician can reliably delineate the OAR contours on CT or MRI (9). Differences in delineation among multiple observers results in dosimetric variability, and reports exist of interobserver dosimetric variability in intracavitary brachytherapy on CT (15) and in hybrid intracavitary/interstitial brachytherapy on MRI (16, 17). Given the proximity between interstitial needles and OARs, and the higher number of needles and different needle arrangement used in interstitial brachytherapy compared with hybrid intracavitary/interstitial brachytherapy, the results reported in the literature may differ from the ones in template-based interstitial brachytherapy.

In this study we quantify the geometric and dosimetric interobserver variability in CT-guided gynecologic interstitial implantation using a 14-patient series retrospectively contoured by 13 physicians.

Methods and Materials

Records of 14 patients treated with gynecologic high-dose-rate interstitial brachytherapy were retrospectively analyzed as part of an institutional review board–approved protocol. One patient (patient 5) did not have a rectum owing to a previous history of rectal cancer and was excluded from the rectal analysis.

Clinical information

All patients received gynecologic interstitial brachytherapy for a gynecologic cancer extending into the vagina. Diagnoses were recurrent endometrial cancer (n = 8), recurrent vulvar cancer (n = 1), cervical cancer (n = 3), vaginal cancer (n = 1) and uterine cancer (n = 1).

All implantations were performed using plastic equipment (interstitial needles, Syed-Neblett template and vaginal obturator) under image guidance, either CT (n = 11) or 3-Tesla MR followed by a CT scan immediately after implantation (n = 3). All CT scans were performed with CT-compatible copper dummy markers inserted in the catheters. Optimization was performed using either a manual or a mixed manual and graphic optimization technique with the goal of achieving the following constraints on the EQD2 (equivalent dose in 2 Gy per fraction) sum of brachytherapy and external beam dose: CTV D₉₀ between 70 Gy and 80 Gy, bladder D_{2cc} <90-95 Gy, rectum D_{2cc} <70-75 Gy and sigmoid D_{2cc} <70-75 Gy. Other clinical information is reported in Table e1 (available online).

Contouring

Thirteen physicians (6 attending radiation oncologists practicing gynecologic brachytherapy in 5 different institutions, and 8 residents in radiation oncology who have rotated through a gynecologic brachytherapy service) independently contoured rectum, bladder and sigmoid on 14 CT scans associated with the 14 patients in this study. All scans were performed 2 days after insertion with the implant still in place ("second-day scan"), according to a protocol previously described (18). The dosimetry was based on the dwell loadings of the clinical plan applied to the catheter geometry of the second-day scan.

The physicians agreed on guidelines for the delineation of the rectum, bladder and sigmoid before contouring based on Radiation Therapy Oncology Group contouring guidelines (19).

Each physician did not have access to the other observers' contour sets and the contouring was performed without any time constraints.

Interobserver geometric variability

The geometric relationship between the contours delineated by the physicians was described

by the generalized conformity index ($CI_{gen} = \frac{\sum_{i,j} V_i \cap V_j}{\sum_{i,j} V_i \cup V_j}$), that is the ratio between the sum of all the intersection between the volumes contoured by any given i,j pair of physicians and the sum of the union between the volumes contoured by any given i,j pair of physicians

(20). Generalized κ ($= \frac{p_0 - p_e}{1 - p_e}$, where p_0 is the observed agreement between observers, and p_e is the expected agreement based on chance alone) and the associated standard error

($= \sqrt{\frac{p_0(1 - p_0)}{n(1 - p_e)^2}}$, where n is the total number of observations) were calculated to assess contour agreement versus agreement occurring by chance alone (21-23). (See online supplemental material for additional details.) Standard nomenclature for the agreement levels associated with the κ values was used (21): poor ($\kappa < 0$), slight (κ 0.01-0.2), fair (κ

0.21-0.3), moderate (κ 0.41-0.6), substantial (κ 0.61-0.8), and almost perfect (κ 0.81-1). The means across physicians of the OAR volumes (in cm³) and their associated coefficient of variation (CV = standard deviation/mean) were calculated.

Interobserver dosimetric variability

The clinical plan associated with each patient was used to calculate the $D_{0.1cc}$ and the D_{2cc} for each contour set and for each physician. The mean across physicians and the associated CV were calculated. The $D_{0.1cc}$ and D_{2cc} CV quantifies the interobserver variability in the single fraction associated with the CT scan used for contouring. We also calculated the standard deviation of the total EQD2 brachytherapy dose (SD^{TOT}), calculated with $\alpha/\beta = 3$ and assuming all fraction metrics to be equal to second-day scan metrics.

Statistical analysis

Kappa statistics and CI_{gen} were calculated using MATLAB (The MathWorks, Natick, MA).

Statistical significance of the difference between 2 metrics was assessed with a 2-tailed paired Student *t* test. Significance was associated with a *P* value <.05.

Results

Patient 5, who did not have a rectum, was excluded from the rectum data analysis. Patient 13, who did not have a sigmoid in the field of view of the scan, was excluded from the sigmoid data analysis.

All results are reported as the mean across the patient population $\pm$ 1 standard deviation calculated across the patient population.

Interobserver geometric variability

The κ agreement across physicians was substantial or almost perfect for all patients for bladder (mean κ of 0.77 ± 0.06 ; 9 of 14 patients with substantial agreement, 5 with almost perfect agreement) and rectum (mean κ of 0.74 ± 0.06 ; 12 of 13 patients with substantial agreement, 1 with almost perfect agreement). The κ agreement across physicians was as low as slight for some patients for sigmoid (mean κ of 0.33 ± 0.20 ; 4 of 13 patients with slight, 4 with fair, 4 with moderate, and 1 with substantial agreement). The κ for sigmoid was significantly lower than those for bladder ($P < .001$) and rectum ($P < .001$). No statistically significant difference was observed between bladder and rectum κ statistics ($P = .19$).

Bladder and rectum showed a similar mean CI_{gen} (0.70 ± 0.08 for bladder, 0.67 ± 0.08 for rectum; $P = .34$) and volume CV ($20\% \pm 6\%$ for bladder, $20\% \pm 5\%$ for rectum; $P = .74$). Sigmoid mean CI_{gen} was 0.26 ± 0.17 (volume CV was $82\% \pm 42\%$), which was significantly lower than those for bladder and rectum ($P = .001$).

Examples of bladder, rectum and sigmoid delineations are presented in Figures 1 and 2. Kappa values, conformity indices and organ volume sizes for all patients are reported in Tables 1-3.

Interobserver dosimetric variability

Dosimetric variability was as follows: for bladder, $CV = 31\% \pm 19\%$ ($SD^{TOT} = 72 \pm 64$ Gy) for $D_{0.1cc}$ and $CV = 16\% \pm 10\%$ ($SD^{TOT} = 9 \pm 6$ Gy) for D_{2cc} ; for rectum, $CV = 11\% \pm 5\%$ ($SD^{TOT} = 16 \pm 17$ Gy) for $D_{0.1cc}$ and $CV = 7\% \pm 2\%$ ($SD^{TOT} = 4 \pm 3$ Gy) for D_{2cc} ; and for sigmoid, $CV = 39\% \pm 28\%$ ($SD^{TOT} = 12 \pm 18$ Gy) for $D_{0.1cc}$ and $CV = 34\% \pm 19\%$ ($SD^{TOT} = 4 \pm 4$ Gy) for D_{2cc} . Despite the similar geometric agreement, rectum metrics showed significantly less interobserver variability (CV and SD^{TOT}) than those for bladder ($P = .004$). No significant difference was observed between sigmoid and bladder CV $D_{0.1cc}$ ($P = .27$), whereas sigmoid showed a significantly higher CV for all other metrics ($P = .004$). Because of the lower sigmoid doses compared with bladder and rectum doses in our series, this higher percent variability did not correlate with a higher variability on total EQD2 brachytherapy dose (SD^{TOT}). Sigmoid SD^{TOT} was significantly lower ($P < .05$) for all metrics except for rectum D_{2cc} ($P = .85$). All dosimetric variability data are summarized in Table 1 for bladder, Table 2 for rectum and Table 3 for sigmoid. Mean values of all metrics are reported in Table 4.

In some cases the OAR D_{2cc} values reported in this study would result in a higher total D_{2cc} dose than our customary constraints would allow. This is due to the target volume abutting the OAR and to the expansion of rectum and bladder between the CT scan at the time of planning and the second-day CT scan used in this study. This effect was previously described (18).

Discussion

We found substantial to almost perfect geometric agreement across 13 physicians in the delineation of the bladder and the rectum, with no statistically significant difference in the agreement level of bladder and rectum. We found an overall fair agreement across the physicians in the delineation of the sigmoid, with results ranging from slight to substantial agreement. The worse agreement of the sigmoid delineations compared with the bladder and rectum delineations was statistically significant. This result is likely due to uncertainty in the inferior–superior extent of the sigmoid and possible uncertainty regarding the assignment of the visible colon to the sigmoid or to other parts of the bowel. We found that the interobserver dosimetric variability of the D_{2cc} metric for a single fraction (CV) was 16% for bladder, 7% for rectum and 34% for sigmoid. Higher CVs were observed for $D_{0.1cc}$ (bladder = 31%, rectum = 11%, sigmoid = 39%). The interobserver variability measured as the standard deviation across observers of the dose for $D_{0.1cc}$ (bladder = 72 Gy, rectum = 16 Gy, sigmoid = 12 Gy) and D_{2cc} (bladder = 9 Gy, rectum = 4 Gy, sigmoid = 4 Gy) shows that sigmoid uncertainties $>30\%$ did not correlate in our series with a meaningful total EQD2 brachytherapy dose variability. This is explained by the lower mean total EQD2 brachytherapy dose to the sigmoid ($D_{0.1cc} = 14$ Gy, $D_{2cc} = 7$ Gy) in our series compared with bladder ($D_{0.1cc} = 120$ Gy, $D_{2cc} = 40$ Gy) and rectum ($D_{0.1cc} = 74$ Gy, $D_{2cc} = 32$ Gy). The rectum dosimetric variability was significantly lower than the bladder dosimetric variability despite presenting an almost identical geometric variability. Our results suggest that minor discrepancies in the delineation of the posterior bladder wall, where high-dose

regions may be present, may correlate with a meaningful dosimetric variability, especially for the $D_{0.1cc}$ metric.

The substantial agreement in the delineation of rectum and bladder confirms that the general shape of rectum and bladder is easily identifiable on CT scans, but uncertainties may remain in the fine delineation of the rectum and bladder interfaces with the surrounding tissues. The same qualitative argument was first proposed in a study by Saarnak et al (15) in an analysis of the variability of bladder and rectum D_{2cc} across 3 physicians for 10 cervical cancer patients treated with CT-guided low-dose-rate tandem-and-ovoids brachytherapy. Sigmoid delineation agreement was very inconsistent in our series, with less than moderate agreement across physicians for 8 of 13 patients. A recent study of interobserver variability in OAR delineation by Hellebust et al (16), in which 10 physicians contoured on MRI the bladder, rectum and sigmoid of 6 cervical cancer patients treated with a tandem and ring with interstitial needles implanted through holes in the ring, suggests that the uncertain classification of the interface between sigmoid and rectum may account for the poor interobserver agreement.

The literature has few quantitative reports of conformity index of OAR delineations on CT and none for interstitial brachytherapy. There is one qualitative description of the interobserver difference in delineation for bladder and rectum (15). Although Hellebust et al (16) do not report the conformity index and κ for MRI-delineated OARs, the variability in organ volume is reported. An analysis of the volume data shows a volume CV of $11\% \pm 4\%$ for bladder, $21\% \pm 11\%$ for rectum and $40\% \pm 3\%$ for sigmoid. Using a similar analysis on the data reported by Saarnak et al (15) for organ delineation on CT, the volume CV (also called the SD%) was $9\% \pm 6\%$ for bladder and $26\% \pm 22\%$ for rectum (sigmoid data were not reported). Our results depart from the existing literature; in our set, the geometric variability of rectum and bladder was comparable and not statistically significant (with similar volume CV of 20%, $P = .74$), whereas the existing literature finds a lower variability for bladder than for rectum. This discrepancy may be due to scanning techniques (improvements in CT scanning technology since Saarnak et al [15], MRI vs CT for Hellebust et al [16]) or to differences in guidelines for the contouring of the organs. Our results also show a higher geometric variability for the sigmoid (volume CV of $82\% \pm 42\%$) compared with MRI delineation (16). The increased visibility of the bowel on MR compared with CT may explain the lower variability in sigmoid contouring, because the allocation of structures between rectum, bowel and sigmoid may be facilitated.

We found a mean bladder interobserver dosimetric variability CV of 31% for $D_{0.1cc}$ and of 16% for D_{2cc} . The results for D_{2cc} are in line with the findings in Saarnak et al (15), which showed a mean CV of 10% for bladder D_{2cc} on CT ($D_{0.1cc}$ was not reported). Hellebust et al (16) report a CV (called SD% in that article) of 5.4% for bladder D_{2cc} on MRI, which suggests that MRI may provide a meaningful increase in visibility of the interface between the bladder wall and surrounding tissue. The presence of dosimetric interobserver variability despite the good overall geometric delineation agreement is likely due to relatively small differences in delineation of the posterior bladder wall in regions of high dose. The bladder $D_{0.1cc}$ metric is not used during plan optimization in our clinic but it has been suggested that it should be reported (11). Our results show a variability in the total EQD2 brachytherapy

$D_{0.1cc}$ for bladder of 72 Gy, suggesting that the use and reporting of the bladder $D_{0.1cc}$ for interstitial brachytherapy should be carefully weighed against the high interobserver variability associated with it.

We found a mean rectum interobserver dosimetric variability CV of 11% for $D_{0.1cc}$ and of 7% for D_{2cc} . The variability for CV D_{2cc} is lower than what was previously reported by Saarnak et al (15) (CV of 11%) and by Hellebust et al (16) (CV of 8%). The lower variability in rectal dose observed in our series compared with the literature, where the main (16) or only (15) contributing source of dose gradient was from an intracavitary applicator, may be explained by a more conservative approach to rectal sparing afforded by the presence of many catheters in the clinical target volume in template-based interstitial implantations, allowing minimization of the dose gradient around the rectum. Such an approach is routine in our clinic to reduce the plan sensitivity to changes in rectal filling during the course of a multi-fractionated treatment. Similarly, the lower variability associated with the rectum compared with bladder was not observed in the literature (15, 16) and is not justified by a more accurate delineation of the rectum than the bladder, given the similar geometric variability for the 2 structures. The planning strategy adopted in our clinic to minimize high-dose regions around the rectum, while allowing some high-dose regions close to the anterior bladder wall owing to its higher tolerance to radiation, may explain the observed difference in results.

We found a mean sigmoid interobserver dosimetric variability CV of 39% for $D_{0.1cc}$ and of 34% for D_{2cc} . The variability for D_{2cc} is higher than what was previously reported by Hellebust et al (16) (SD% or CV of 11%) which is expected given the higher geometric variability. Although the sigmoid CV is significantly higher than the bladder and rectum CV ($P < .001$), the observed total EQD2 brachytherapy dose variability of 12 Gy for $D_{0.1cc}$ and 4 Gy for D_{2cc} is equal to or smaller than the variability associated with the rectum and bladder, owing to the lower mean total sigmoid dose in our patient population. This effect has not been reported for intracavitary applications, possibly owing to a higher dose to the sigmoid related to potential close proximity between sigmoid and tandem (24).

No consensus exists in the literature as to the best variable to describe interobserver variability. We presented in this work the CV to describe a distribution in a single number, the conformity index (CI_{gen}) to describe the overlap between contours and κ to measure statistical agreement. Fotina et al (25) suggest a similar 3-group separation of parameters according to their associated methodology. Other choices are possible; for example the intraclass correlation coefficient has been suggested for statistical analysis to evaluate how many contouring physicians may be needed (25). The analysis of region of overlap can be performed with topographic tools; this was not performed in this work given the amount of information already reported and will be considered for future analysis. Because the areas of dosimetric interest are 0.1-cm³ and 2-cm³ volumes located near high-dose regions, a parameter providing a measurement of agreement for given regions within an OAR may provide quantitative insight into the correlation between geometric and dosimetric variability.

Conclusion

A substantial rectum and bladder geometric agreement (mean $\kappa > 0.74$) was associated with dosimetric interobserver variability of 11% for rectum $D_{0.1cc}$, 7% for rectum D_{2cc} and 16% for bladder D_{2cc} . A larger dosimetric variability (31%) was observed for the bladder $D_{0.1cc}$, probably owing to sharp dose gradients by the posterior bladder wall.

A fair sigmoid geometric agreement (mean $\kappa = 0.33$) and dosimetric interobserver variability ($>30\%$ for $D_{0.1cc}$ and D_{2cc}) for the sigmoid were associated with a small total dose variability (12 Gy for $D_{0.1cc}$, 4 Gy for D_{2cc}), owing to low dose and low dose gradients in the sigmoid region for the patients in our series. Contouring variability is an important factor in the comprehensive assessment of pelvic normal tissue dosimetry.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Barbara Silver for reviewing the manuscript.

This work was partially funded by National Institutes of Health grant R21 CA167800 (Principal Investigator, A.N.V.).

References

1. Abbe R. The use of radium in malignant disease. *Lancet*. 1913; 2:524–527.
2. Bloedorn FG. Application of the Paterson-Parker system in interstitial radium therapy. *Am J Roentgenol Radium Ther Nucl Med*. 1956; 75:457–475.
3. Quimby EH, Castro V. The calculation of dosage in interstitial radium therapy. *Am J Roentgenol Radium Ther Nucl Med*. 1953; 70:739–749.
4. Bellotti JE, Kagan AR, Wollin M, et al. Application of the ICRU Report 38 reference volume concept to the radiotherapeutic management of recurrent endometrial and cervical carcinoma. *Radiother Oncol*. 1993; 26:254–259. [PubMed: 8316655]
5. Erickson B, Albano K, Gillin M. CT-guided interstitial implantation of gynecologic malignancies. *Int J Radiat Oncol Biol Phys*. 1996; 36:699–709. [PubMed: 8948356]
6. Lee LJ, Damato AL, Viswanathan AN. Clinical outcomes of high-dose-rate interstitial gynecologic brachytherapy using real-time CT guidance. *Brachytherapy*. 2013; 12:303–310. [PubMed: 23491023]
7. Viswanathan AN, Cormack R, Holloway CL, et al. Magnetic resonance-guided interstitial therapy for vaginal recurrence of endometrial cancer. *Int J Radiat Oncol Biol Phys*. 2006; 66:91–99. [PubMed: 16839709]
8. Kapur T, Egger J, Damato A, et al. 3-T MR-guided brachytherapy for gynecologic malignancies. *Magn Reson Imaging*. 2012; 30:1279–1290. [PubMed: 22898699]
9. Beriwal S, Demanes DJ, Erickson B, et al. American Brachytherapy Society consensus guidelines for interstitial brachytherapy for vaginal cancer. *Brachytherapy*. 2012; 11:68–75. [PubMed: 22265440]
10. Potter R, Georg P, Dimopoulos JC, et al. Clinical outcome of protocol based image (MRI) guided adaptive brachytherapy combined with 3D conformal radiotherapy with or without chemotherapy in patients with locally advanced cervical cancer. *Radiother Oncol*. 2011; 100:116–123. [PubMed: 21821305]

11. Potter R, Haie-Meder C, Van Limbergen E, et al. Recommendations from gynaecological (GYN) GEC ESTRO working group (II): concepts and terms in 3D image-based treatment planning in cervix cancer brachytherapy-3D dose volume parameters and aspects of 3D image-based anatomy, radiation physics, radiobiology. *Radiother Oncol.* 2006; 78:67–77. [PubMed: 16403584]
12. Viswanathan AN, Erickson BA. Three-dimensional imaging in gynecologic brachytherapy: A survey of the American Brachytherapy Society. *Int J Radiat Oncol Biol Phys.* 2010; 76:104–109. [PubMed: 19619956]
13. Georg P, Lang S, Dimopoulos JC, et al. Dose-volume histogram parameters and late side effects in magnetic resonance image-guided adaptive cervical cancer brachytherapy. *Int J Radiat Oncol Biol Phys.* 2011; 79:356–362. [PubMed: 20385450]
14. Lee LJ, Viswanathan AN. Predictors of toxicity after image-guided high-dose-rate interstitial brachytherapy for gynecologic cancer. *Int J Radiat Oncol Biol Phys.* 2012; 84:1192–1197. [PubMed: 22592049]
15. Saarnak AE, Boersma M, van Bunningen BN, et al. Inter-observer variation in delineation of bladder and rectum contours for brachytherapy of cervical cancer. *Radiother Oncol.* 2000; 56:37–42. [PubMed: 10869753]
16. Hellebust TP, Tanderup K, Lervag C, et al. Dosimetric impact of interobserver variability in MRI-based delineation for cervical cancer brachytherapy. *Radiother Oncol.* 2013; 107:13–19. [PubMed: 23462706]
17. Petric P, Hudej R, Rogelj P, et al. Uncertainties of target volume delineation in MRI guided adaptive brachytherapy of cervix cancer: A multi-institutional study. *Radiother Oncol.* 2013; 107:6–12. [PubMed: 23453539]
18. Damato AL, Cormack RA, Viswanathan AN. Characterization of implant displacement and deformation in gynecologic interstitial brachytherapy. *Brachytherapy.* 2014; 13:100–109. [PubMed: 24269145]
19. Gay HA, Barthold HJ, O'Meara E, et al. Pelvic normal tissue contouring guidelines for radiation therapy: A Radiation Therapy Oncology Group consensus panel atlas. *Int J Radiat Oncol Biol Phys.* 2012; 83:e353–e362. [PubMed: 22483697]
20. Kouwenhoven E, Giezen M, Struikmans H. Measuring the similarity of target volume delineations independent of the number of observers. *Phys Med Biol.* 2009; 54:2863–2873. [PubMed: 19384002]
21. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics.* 1977; 33:363–374. [PubMed: 884196]
22. Fleiss, JL. *Statistical Methods for Rates and Proportions.* 2. New York: John Wiley & Sons; 1981. p. 38–46.
23. Allozi R, Li XA, White J, et al. Tools for consensus analysis of experts' contours for radiotherapy structure definitions. *Radiother Oncol.* 2010; 97:572–578. [PubMed: 20708285]
24. Holloway CL, Racine ML, Cormack RA, et al. Sigmoid dose using 3D imaging in cervical-cancer brachytherapy. *Radiother Oncol.* 2009; 93:307–310. [PubMed: 19665244]
25. Fotina I, Lutendorf-Caucig C, Stock M, et al. Critical discussion of evaluation parameters for inter-observer variability in target definition for radiation therapy. *Strahlenther Onkol.* 2012; 188:160–167. [PubMed: 22281878]

Summary

Thirteen physicians contoured rectum, bladder, and sigmoid on 14 patients' CT scans. All patients received image-guided interstitial gynecologic brachytherapy. Substantial rectum and bladder (mean $\kappa > 0.74$) and fair sigmoid (mean $\kappa = 0.33$) geometric agreement was observed. The coefficient of variation (CV) for $D_{0.1cc}$ for bladder and $D_{0.1cc}$ and D_{2cc} sigmoid was $> 30\%$. The CV for all other $D_{0.1cc}$ and D_{2cc} was 16%.

Fig. 1.
Example of contouring variability: patient 10.

Fig. 2.
Example of contouring variability: patient 3.

Table 1

Bladder geometric and dosimetric variability data

Patient no.	κ	κ SE	CI _{gen}	Volume $\pm$ SD (cm ³)	Volume CV (%)	D _{0,rec} CV (%)	D _{0,rec} SD ^{TOT} (Gy)	D _{2ec} CV (%)	D _{2ec} SD ^{TOT} (Gy)
1	0.81	<0.001	0.71	85 $\pm$ 13	15	57	150	29	21
2	0.76	<0.001	0.65	44 $\pm$ 8	18	12	8	12	4
3	0.84	<0.001	0.77	49 $\pm$ 7	14	16	15	9	3
4	0.74	<0.001	0.65	55 $\pm$ 11	20	14	23	9	7
5	0.72	<0.001	0.63	85 $\pm$ 23	27	16	8	9	2
6	0.74	<0.001	0.67	85 $\pm$ 24	28	19	14	13	4
7	0.68	<0.001	0.60	67 $\pm$ 19	28	22	52	12	6
8	0.80	<0.001	0.73	111 $\pm$ 27	24	57	105	26	17
9	0.83	<0.001	0.78	86 $\pm$ 15	18	27	127	10	13
10	0.78	<0.001	0.70	67 $\pm$ 12	17	20	92	9	14
11	0.71	<0.001	0.62	73 $\pm$ 17	23	49	87	24	7
12	0.82	<0.001	0.76	91 $\pm$ 14	15	19	15	9	3
13	0.87	<0.001	0.86	109 $\pm$ 8	7	68	216	39	16
14	0.69	<0.001	0.60	79 $\pm$ 21	26	34	96	12	13
Average across patients	0.77	<0.001	0.70	78	20	31	72	16	9
SD across patients	0.06	<0.001	0.08	20	6	19	64	10	6
P value vs rectum	.193	-	.336	.206	.740	.003	.004	.002	.002
P value vs sigmoid	<.001	-	<.001	.001	<.001	.270	.008	.004	.020

Abbreviations: CI_{gen} = generalized conformity index; CV = ratio of the standard deviation to the mean; SD = standard deviation; SD^{TOT} = standard deviation of the total EQD2 (equivalent dose in 2 Gy per fraction) brachytherapy dose; SE = standard error for κ .

Table 2

Rectum geometric and dosimetric variability data

Patient no.	κ	κ SE	CI_{gen}	Volume $\pm$ SD (cm ³)	Volume CV (%)	$D_{0,rec}$ CV (%)	$D_{0,rec}$ SD ^{TOT} (Gy)	D_{2ec} CV (%)	D_{2ec} SD ^{TOT} (Gy)
1	0.70	<0.001	0.63	55 $\pm$ 8	15	14	9	8	3
2	0.73	<0.001	0.66	34 $\pm$ 6	17	6	3	3	1
3	0.80	<0.001	0.75	62 $\pm$ 11	17	15	8	8	2
4	0.79	<0.001	0.73	75 $\pm$ 11	15	12	16	5	3
5	-	-	-	-	-	-	-	-	-
6	0.70	<0.001	0.59	98 $\pm$ 22	23	14	17	5	2
7	0.63	<0.001	0.54	32 $\pm$ 9	29	6	2	5	1
8	0.73	<0.001	0.65	67 $\pm$ 14	21	11	9	7	3
9	0.67	<0.001	0.59	62 $\pm$ 16	26	25	56	8	8
10	0.72	<0.001	0.65	40 $\pm$ 9	23	6	6	5	3
11	0.79	<0.001	0.70	120 $\pm$ 30	25	10	45	12	6
12	0.76	<0.001	0.67	56 $\pm$ 11	19	11	6	7	2
13	0.77	<0.001	0.72	24 $\pm$ 6	24	8	3	9	2
14	0.86	<0.001	0.83	105 $\pm$ 10	10	9	24	7	11
Average across patients	0.74	<0.001	0.67	0	20	11	16	7	4
SD across patients	0.06	<0.001	0.08	0	5	5	17	2	3
P value vs bladder	.193	-	.336	.206	.740	.003	.004	.002	.002
P value vs sigmoid	<.001	-	<.001	.014	.00	.002	.044	.000	.854

Abbreviations as in Table 1.

Patient 5 did not have a rectum.

Table 3

Sigmoid geometric and dosimetric variability data

Patient no.	κ	κ SE	CI _{gen}	Volume $\pm$ SD (cm ³)	Volume CV (%)	D _{0,1cc} CV (%)	D _{0,1cc} SD ^{TOT} (Gy)	D _{2cc} CV (%)	D _{2cc} SD ^{TOT} (Gy)
1	0.31	0.00	0.22	24 $\pm$ 20	82	52	7	53	5
2	0.12	0.00	0.08	6 $\pm$ 5	96	62	8	57	4
3	0.34	0.00	0.28	68 $\pm$ 63	93	40	4	43	2
4	0.57	0.00	0.44	48 $\pm$ 17	35	27	4	13	1
5	0.07	0.00	0.06	17 $\pm$ 17	101	105	66	59	9
6	0.46	<0.001	0.40	106 $\pm$ 63	59	24	5	20	3
7	0.51	0.00	0.37	23 $\pm$ 9	39	24	2	19	1
8	0.69	<0.001	0.63	41 $\pm$ 12	29	5	1	4	0
9	0.10	0.00	0.07	27 $\pm$ 35	131	71	33	54	11
10	0.33	0.00	0.26	32 $\pm$ 31	95	14	8	12	3
11	0.26	0.00	0.19	36 $\pm$ 22	62	44	1	50	1
12	0.42	0.00	0.31	64 $\pm$ 43	66	30	3	31	2
13	-	-	-	-	-	-	-	-	-
14	0.10	0.00	0.08	32 $\pm$ 57	180	11	18	28	9
Average across patients	0.33	0.00	0.26	8	82	39	12	34	4
SD across patients	0.20	<0.001	0.17	4	42	28	18	19	4
P value vs. bladder	<.001	-	<.001	.001	<.001	.270	.008	.004	.020
P value vs. rectum	<.001	-	<.001	.014	.00	.002	.044	.000	.854

Abbreviations as in Table 1.

Patient #13 did not have a sigmoid in the field-of-view.

Table 4

Mean values of the single fraction $D_{0.1cc}$ and D_{2cc} (not in EQD2 [equivalent dose in 2 Gy per fraction]) and of the total brachytherapy dose calculated with the EQD2 formalism

Patient no.	Bladder				Rectum				Sigmoid			
	$D_{0.1cc}$	Total EQD2	$D_{0.1cc}$	D_{2cc}	Total EQD2	$D_{0.1cc}$	D_{2cc}	Total EQD2	$D_{0.1cc}$	D_{2cc}	Total EQD2	D_{2cc}
1	7.2	161.7	46.3	3.7	38.8	2.4	23.8	1.2	10.1	0.9	6.7	6.7
2	3.7	44.8	23.3	2.4	30.1	2.1	18.8	1.2	8.7	0.8	3.8	3.8
3	6.0	54.3	22.4	3.5	32.7	2.7	15.3	1.5	5.7	1.0	3.6	3.6
4	8.3	95.8	45.9	7.1	72.0	5.1	41.0	1.9	9.4	1.3	5.4	5.4
5	3.7	30.3	17.3	2.6	-	-	-	3.7	31.0	1.8	6.9	6.9
6	4.2	43.0	23.4	2.8	74.1	3.0	25.4	1.9	12.4	1.3	7.5	7.5
7	9.0	155.2	34.0	3.6	18.2	1.9	12.7	1.0	6.0	0.8	4.0	4.0
8	5.6	80.2	35.2	3.7	51.4	3.4	29.9	2.0	14.3	1.5	9.3	9.3
9	14.3	259.7	81.1	7.6	109.9	6.0	55.0	3.5	22.1	2.3	10.4	10.4
10	10.6	264.9	95.1	4.7	66.0	3.5	41.6	3.5	41.7	2.3	21.5	21.5
11	7.3	103.1	18.6	2.7	255.3	3.8	31.2	0.5	1.9	0.3	1.1	1.1
12	4.2	43.8	22.9	2.8	36.4	2.4	18.5	1.2	7.4	0.9	5.0	5.0
13	10.0	207.2	25.1	3.2	24.4	2.5	16.6	-	-	-	-	-
14	10.1	142.6	65.4	6.7	152.1	7.9	87.0	4.5	15.5	2.7	7.2	7.2
Average across patients	7.4	120.5	39.7	4.0	73.9	3.6	32.1	2.1	14.3	1.4	7.1	7.1
SD across patients	3.1	80.3	24.6	1.7	66.3	1.8	20.6	1.2	11.2	0.7	5.0	5.0
P value vs bladder	-	-	-	-	.076	.081	.096	<.001	<.001	<.001	<.001	<.001
P value vs rectum	.083	.076	.096	.081	-	-	-	.001	.007	<.001	.001	.001
P value vs sigmoid	<.001	<.001	<.001	<.001	.007	<.001	.001	-	-	-	-	-

Abbreviations as in Table 1.

All values are grays (Gy) except where noted. Patient #5 did not have a rectum. Patient #13 did not have a sigmoid in the field of view.