UC San Diego
UC San Diego Previously Published Works

Title

Evaluation of dose differences between intracavitary applicators for cervical brachytherapy using knowledge-based models

Permalink

https://escholarship.org/uc/item/9vn4n9gt

Journal

Brachytherapy, 20(6)

ISSN

1538-4721

Authors

Kallis, Karoline
Mayadev, Jyoti
Covele, Brent
et al.

Publication Date

2021-11-01

DOI

10.1016/j.brachy.2021.08.010

Supplemental Material

https://escholarship.org/uc/item/9vn4n9gt#supplemental

Copyright Information

Peer reviewed

1 **Evaluation of dose differences between intracavitary applicators for cervical brachytherapy using knowledge-based models**

4

5 Karoline Kallis , Jyoti Mayadev , Brent Covele , Derek Brown , Daniel Scanderbeg , Aaron Simon ,
6 Helena Frisbie-Firsching , Catheryn M. Yashar , John P. Einck , Loren K. Mell , Kevin L. Moore , Sandra
7 M. Meyers

8 Department of Radiation Medicine & Applied Sciences, University of California San Diego, San Diego, CA

9

10

11

Purpose: Currently, there is a lack of patient-specific tools to guide brachytherapy planning and applicator choice for cervical cancer. The purpose of this study is to evaluate the accuracy of organ-at-risk (OAR) dose predictions using knowledge-based intracavitary models, and the use of these models and clinical data to determine the dosimetric differences of tandem-and-ring (T&R) and tandem-and-ovoids (T&O) applicators.

Materials and Methods: Knowledge-based models, which predict organ D_{2cc}, were trained on 77/75 cases and validated on 32/38 for T&R/T&O applicators. Model performance was quantified using $\Delta D_{2cc} = D_{2cc,actual} - D_{2cc,predicted}$, with standard deviation ($\sigma(\Delta D_{2cc})$) representing precision. Model-predicted applicator dose differences were determined by applying T&O models to T&R cases, and vice versa, and compared to clinically-achieved D_{2cc} differences. Applicator differences were assessed using a Student's t-test ($p<0.05$ significant).

Results: Validation T&O/T&R model precision was 0.65/0.55Gy, 0.55/0.38Gy, and 0.43/0.60Gy for bladder, rectum and sigmoid, respectively, and similar to training. When applying T&O/T&R models to T&R/T&O cases, bladder, rectum and sigmoid D_{2cc} values in EQD2 were on average 5.69/2.62Gy, 7.31/6.15Gy and 3.65/0.69Gy lower for T&R, with similar HRCTV volume and coverage. Clinical data also showed lower T&R OAR doses, with mean EQD2 D_{2cc} deviations of 0.61Gy, 7.96Gy ($p<0.01$) and 5.86Gy ($p<0.01$) for bladder, rectum and sigmoid.

Conclusion: Accurate knowledge-based dose prediction models were developed for two common intracavitary applicators. These models could be beneficial for standardizing and improving the quality of brachytherapy plans. Both models and clinical data suggest that significant OAR sparing can be achieved with T&R over T&O applicators, particularly for the rectum.

Introduction

Brachytherapy is an important component of the standard of care treatment for cervical cancer, typically used alongside external beam radiation therapy (EBRT) and chemotherapy. The therapy is linked to increased survival rates (1–3) and allows a dose escalation in high-risk regions with optimal sparing of organs-at-risk (OAR) (4). Brachytherapy is commonly administered using tandem and ovoids (T&O) or tandem and ring (T&R) applicators. Given that the source is guided by the chosen applicator, the achievable dosimetry is largely dictated by the implanted applicator. These applicators are often used interchangeably, although there are proven dosimetric differences (5–10), such as decreased bladder, rectum and sigmoid D_{2cc} values using the T&R over the T&O applicator.

Population-based protocols such as EMBRACE provide recommendations for dose planning objectives and guidance on needle supplementation (11, 12). However, there are very few tools available to guide and standardize dose optimization and applicator choice for individual patients, and most decisions depend on the physician's preference and experience. Knowledge-based models have proved to be helpful in EBRT for plan quality control, standardization and automated high quality treatment planning (13–20). These models use data from prior patient treatments to make dose predictions for new patients based on anatomic and geometric features. But despite their proven benefits, knowledge based models are not as common in brachytherapy (21–23).

A knowledge-based dose prediction model has already been validated for T&O applicators, but the method has not yet been extended to other applicators (23). Knowledge-based models could offer additional value in brachytherapy by providing insight into the achievable dosimetry of different applicators. With dose prediction models for both applicators, dosimetric differences between T&O and T&R applicators can be investigated using techniques beyond simple comparisons of clinical data between patient cohorts. This could help to inform applicator choice in the clinic, leading to standardized decision-making with less reliance on physician preference. The purpose of this study was to evaluate the

65 accuracy of OAR dose prediction using knowledge-based intracavitary models for both T&R and T&O

66 applicators, and to use these models and clinical data to determine dosimetric differences of the two

67 applicators.

68 **Materials and Methods**

69 *Patient Cohort*

70   Cervical cancer patients receiving brachytherapy treatment using either a T&R or T&O applicator

71 were included in the study (UCSD IRB #200065C). Treatment fractions that featured additional

72 implanted needles were excluded. In total, 113 T&R and 109 T&O treatment fractions from 55 and 36

73 patients, respectively, were available for the study. The patients were treated with computed tomography

74 (CT) image guided intracavitary high-dose rate brachytherapy within a period of three years.

75 Prescriptions ranged from 5.5Gy to 8.5Gy per fraction, with 3 to 5 fractions in total. Additional clinical

76 details about the patient cohort are summarized in **Table 1**. Treatment planning was performed according

77 to the guidelines defined by EMBRACE II (11, 12), which include the following hard planning criteria:

78 high-risk clinical target volume (HRCTV) D90>85Gy, bladder D_{2cc}<90Gy, rectum D_{2cc}<75Gy and D_{2cc}

79 sigmoid<75Gy. Soft planning aims (recommended but not required) are bladder D_{2cc}<80Gy, rectum

80 D_{2cc}<65Gy, and sigmoid D_{2cc}<70Gy. All dose values are defined in EQD2, i.e. the biologically equivalent

81 dose in 2-Gy fractions.

82 *Knowledge-Based Dose Prediction Models*

83   A detailed description of the knowledge-based dose prediction algorithm can be found in

84 Yusufaly *et al.* (23), which demonstrates model accuracy for cervical cancer patients treated with T&O.

85 Briefly, the models use target to OAR distance to predict OAR dose-volume histograms (DVHs). The

86 models assume that dose conforms to the HRCTV, and the dose fall-off within a particular OAR is

87 dependent on the distance from the HRCTV. The pre-processing required to train models is shown in

88 **Figure 1**. First, shells are created around the HRCTV (Figure 1B), where the inner 20 shells have a width

89   of 2mm, and the outer 12 shells have a width of 6mm. Then OAR sub-volumes are generated from the

90   overlap of each shell with each OAR (Figure 1C). For each OAR sub-volume, a differential DVH is

91   extracted for each trainings case, and all training cases are averaged to produce dose kernels as a function

92   of distance to the HRCTV. In order to predict a DVH for a new patient, the OAR contours are discretized

93   in the same manner. The resulting DVH of the considered OAR is then the sum of differential DVH

94   kernels, weighted by the volume of each OAR sub-volume (13, 23, 24). Once models are produced,

95   DVHs can be predicted for any new patient using the HRCTV and OAR contours alone.

96   Two different sets of models were constructed for T&O and T&R applicators. The T&O (T&R)

97   models were trained on 77 (75) cases and validated on 32 (38) cases. A "case" is defined as a single

98   fraction of a brachytherapy treatment consisting of 3-5 fractions in total. The T&R (T&O) validation data

99   set consisted of 19 (17) independent cases and 19 (15) cases where other fractions were included in model

100 training. Model training and DVH prediction were performed automatically using in-house extensions

101 embedded into MIM (v7.0.1, MIM Software Inc., Cleveland, OH), such that predicted D_{2cc} values could

102 be obtained for a case in under 10 seconds and used to guide treatment planning.

103 *Data Analysis*

104   Data analysis was performed using automated in-house scripts implemented in MATLAB

105 (R2019b, MathWorks, Inc., Natick, MA). Since OAR D_{2cc} is a common metric used to evaluate the quality

106 of clinical treatment plans, this value was extracted from the predicted DVHs and used to determine the

107 precision of the predictions and to quantify the dose difference between applicators. D_{2cc} values were also

108 extracted from actual clinical DVHs. Model performance was quantified using $\Delta D_{2cc} = D_{2cc,\ actual} - D_{2cc,}$

109 $_{predicted}$. The standard deviation (σ) of ΔD_{2cc} represents the model precision while the mean represents model

110 bias. Correlation between actual and predicted D_{2cc} values was evaluated with Pearson correlation

111 coefficients.

112   In order to estimate dosimetric differences between the two applicators, the T&O model was used

113 to make D_{2cc} predictions for 113 cases treated by T&R, and vice versa. Dose differences between the

114　actual plan, with its chosen applicator, and the alternate knowledge-based prediction indicate the potential

115　advantage or disadvantage of that applicator. Model-predicted applicator dose differences were further

116　compared to differences observed in D_{2cc} values from clinically treated plans. HRCTV volume and

117　coverage metrics (D90 and V100) were also compared between clinical plans treated with each

118　applicator. A Student's t-test with a significance level of 0.05 was used to test for significance of

119　deviations.

120　　　The predicted OAR D_{2cc} values of each case were also transformed to EQD2 with α/β = 3, in

121　order to compare to previously reported dose differences between applicators and account for differences

122　in prescription between patients. Assuming that the patients receive the same D_{2cc} values in all fractions,

123　the calculated brachytherapy EQD2 value of a single brachytherapy fraction was multiplied by the total

124　amount of fractions and then the EQD2 dose of prior EBRT was added. The total EQD2 value is referred

125　to as $D_{2cc, EQD2}$ throughout the manuscript.

126　**Results**

127　*Knowledge-Based Models*

128　Model precision ranged between 0.46Gy to 0.70Gy for the T&O cases and between 0.38Gy to 0.68Gy for

129　the T&R cases for the validation dataset. The precision was similar for training and validation datasets for

130　both models (see **Figure 2**, **Figure 3** and **Supplementary Table 1**). T&R (T&O) model bias, represented

131　by average ΔD_{2cc}, for bladder, rectum and sigmoid was -0.02Gy (-0.14Gy), -0.13Gy (-0.06Gy), and -

132　0.21Gy (-0.01Gy), respectively. A negative average of ΔD_{2cc} indicates higher D_{2cc} predictions in

133　comparison to the actual D_{2cc} values. There was a strong correlation between actual and predicted doses,

134　demonstrated by the high Pearson correlation coefficients (see **Figure 2** and **Figure 3**). Overall, there was

135　good agreement between actual and predicted D_{2cc} values and no prediction accuracy difference between

136　the two applicator models.

137　*Clinical Data*

6

138   Dose metrics for the actual, clinical treatment plans and clinical characteristics for each patient

139   cohort are summarized in **Table 1**, and dose differences between T&R and T&O plans are displayed in

140   **Table 2**. Both patient cohorts featured similar HRCTV volumes of around 19cc on average. D90 of

141   HRCTV was, on average, 6.91% higher for T&R cases. Mean ± standard deviation of D_{2cc} for T&O

142   treatment plans were 4.49±1.02Gy for bladder, 3.27±0.91Gy for rectum and 3.67±0.86Gy for sigmoid.

143   T&R values were 4.97±1.15Gy for bladder, 2.56±0.86Gy for rectum and 3.44±1.25Gy for sigmoid.

144   Average differences in EQD2 D_{2cc} were 0.61Gy, 7.96Gy and 5.82Gy for bladder, rectum and sigmoid,

145   respectively, where a positive value indicates T&O had higher dose. When normalizing each D_{2cc} dose to

146   prescription, the corresponding differences were 1%, 16% and 11%. These differences were significant for

147   both rectum and sigmoid. Comparisons of EQD2 OAR dose and HRCTV dose and volume are shown in

148   **Supplementary Figure 1**.

149   *Predicted Dose Differences*

150   Dose differences between applicators were revealed when applying the T&R model to T&O cases and

151   vice versa (see **Table 2** and **Figure 6**). When the T&O model was applied to T&R cases, predicted OAR

152   doses reported in EQD2 (absolute dose relative to prescription), were found to be 5.55Gy (10%) larger on

153   average over all OARs (range 3.65 – 7.31Gy, 7 – 15%), indicating that the model predicted that the T&O

154   applicator would result in hotter OAR dose. Similarly, when the T&R model was applied to T&O cases,

155   predicted OAR doses were found to be on average 3.15Gy (7%) lower than actual T&O clinical data

156   (range 0.69 – 6.15Gy, 2 – 13%). Both models predicted significant dose sparing for bladder and rectum

157   using the T&R applicator over the T&O applicator.

158   **Discussion**

159   This study explores the use of knowledge-based models to predict organ dose for two common

160   brachytherapy applicators and determine possible dose differences between these applicators. It is

161   clinically relevant for clinicians to use predictive tools for decision-making in addition to experience and

162 brachytherapy skill expertise. Furthermore, knowledge-based dose prediction models have found

163 numerous applications in EBRT, including plan quality control and automated planning (13–20), and a

164 few groups have used prior patient data to inform treatment planning or dose prediction in brachytherapy

165 (21–23, 25, 26). However, until now no study has used knowledge-based models to gain insight into

166 brachytherapy applicator differences. Although there are solutions for automating various aspects of

167 brachytherapy treatment planning, such as applicator reconstruction (25, 27) and inverse optimization

168 (28–32), there are currently no tools to guide gynecological applicator choice. The choice of applicator is

169 not standardized and relies on the physician's preference and expertise, which is particularly challenging

170 for inexperienced physicians. The model predictions presented in this work provide insight into the

171 specific dosimetric advantages of each applicator, which could help physicians make informed decisions

172 based on quantitative metrics.

173   The proposed, simple model predicts OAR D_{2cc} with a precision between 0.38-0.70Gy in a few

174 seconds, using only contours as input. Even though T&R and T&O models were trained on different

175 patient groups, they achieved similar precision (see **Supplementary Table 1**). Both models were trained

176 on around 100 cases and proved to have a similar accuracy to the earlier T&O model presented by

177 Yusufaly *et al*. (23) using 356 cases. They reported a model precision between 0.43-0.61Gy for bladder,

178 rectum and sigmoid using cases treated according to either EMBRACE I or II guidelines. In our study, we

179 limited the patient cohort to patients only treated according to the EMBRACE II guidelines to ensure

180 similarity between the patient groups. Our model biases (0.02-0.21Gy) were much less than the standard

181 deviations (0.38-0.70Gy), although there did appear to be a slight trend of a negative bias for all OAR

182 models meaning that D_{2cc} predictions were higher than the actual D_{2cc}. At this point, we don't have an

183 explanation for the effect; however, when applying the same model training and validation procedure to a

184 dataset four times the size (23), model bias was close to zero. Therefore, we suspect noise from smaller

185 statistics, sporadic case-to-case differences in contouring or patient selection within this smaller sample

186   size likely contributed to the model biases observed in this study. As in EBRT (16, 33) these models

187   could be beneficial for plan quality control by providing patient-specific dose objectives to aim for when

188   planning, leading to greater standardization and quality of treatment plans. Further discussion of the

189   advantages and limitations of this model can be found in (23).

190   *Comparison of T&R and T&O Applicators*

191   T&R and T&O applicators are often used interchangeably in the clinic, though both our models

192   and clinical data suggest that there are substantial dosimetric differences between these applicators. In

193   particular, rectal dose was much lower with T&R, which we suspect was due to the rectal retractor. Both

194   model and clinical results suggest that T&R could provide up to 1Gy per brachytherapy fraction of rectal

195   dose sparing.

196   Several previous studies have retrospectively compared the dosimetry and outcome of T&R and

197   T&O applicators based on clinical data (5–10). Biltekin *et al.* (5) found significant dose sparing in the

198   rectum using T&R over T&O, and bladder, sigmoid and rectum D_{2cc} were, on average, 0.94Gy, 0.59Gy

199   and 1.36Gy lower per brachytherapy fraction for cases treated with T&R. These findings agree with our

200   data for rectum and sigmoid (0.71Gy and 0.32Gy, respectively), although our dose differences were

201   slightly smaller and significant. The slight differences could be explained by their reduced sample size of

202   10 patients (26 cases), compared to our 55 T&R and 36 T&O patients (113 and 109 cases).

203   Ma *et al*. (7) compared the short-term clinical outcome for a total of 52 fractions of 13 patients

204   and dose metrics between applicators and found no significant difference, though T&R D_{2cc} values were

205   0.41Gy, 0.48Gy and 0.68Gy lower per fraction for bladder, rectum and sigmoid, respectively.

206   Gursel *et al.* (10) analyzed dosimetric differences of intracavitary applicators for 20 patients and

207   found significantly lower EQD2 D_{2cc} values for T&R of 3.79Gy and 11.90Gy for bladder and rectum.

208   Another study on the results of the EMBRACE I trial reported EQD2 D_{2cc} reductions of 7.7Gy, 3.3Gy and

209   0.8Gy for bladder, rectum and sigmoid with centers utilizing T&R applicators over T&O (6). While our

210   data showed similar trends, the magnitudes of dose reductions in EQD2 were different (0.61Gy, 7.96Gy

211 and 5.86Gy), which could be caused by a few factors. For one, our patient data was sorted per

212 brachytherapy fraction (since some patients received different applicators over the course of treatment),

213 and as a result we computed an effective EQD2 for each case by assuming that the same D_{2cc} value was

214 delivered for all fractions. In contrast, Serban *et al.* grouped data by centers, which were classified

215 according to the most used applicator and allowed up to 20% of cases to be delivered with interstitial

216 needles. In addition, our patients were treated according to EMBRACE II guidelines, which include more

217 conservative planning aims to guide dose optimization. In summary, all studies agree that there are

218 dosimetric differences between applicators and that OAR dose can be spared using a T&R applicator over

219 T&O applicator.

220   HRCTV coverage is another important consideration when evaluating brachytherapy plan quality,

221 and could confound comparisons of OAR dose between applicators. In our study, both patient cohorts had

222 similar sized HRCTV volumes; however, HRCTV coverage was significantly greater for T&R plans

223 relative to T&O (HRCTV V100 and D90 were 1.46% and 6.91% higher, respectively), which is

224 impressive given the greater OAR sparing with T&R. Three other studies also reported higher HRCTV

225 D90 with T&R, one significant (9.0Gy EQD2 (10)), and the other two insignificant (2.4Gy EQD2 (6) and

226 0.044Gy per brachytherapy fraction (5), on average). Previous results for volume metrics found better

227 coverage for T&O: Serban *et al*. reported V85Gy EQD2 was 17.9cc (about 20%) higher for T&O (6),

228 while another study found significantly larger V95%, 85% and 50% for T&O (9). In contrast, Gursel *et al.*

229 (10) reported significantly higher V100 for T&R (5.54%), which agrees with our findings.

230   There are many other factors that influence applicator choice. For instance, the T&O applicator

231 may allow more degrees of freedom with choosing the desired tandem length (34); in contrast, the fixed

232 geometry of the ring relative to the tandem leads provides less flexibility, although dose distributions are

233 more reproducible (9). Because the T&O can often treat further into the uterus, it may be preferred by

234 some physicians for certain patients. In our data, the prescription dose extended more superior for many

235 T&O treatment plans, and thus the higher sigmoid dose was expected for T&O and not necessarily a

236   detriment. Because T&O does not include a rectal retractor, manual vaginal packing is required, which

237   can result in greater variability between treatments, patient-specific anatomy and physician skillset. We

238   have not explored the difference in total treatment time, but other studies have reported longer treatment

239   times with T&O applicators (5, 9).

240   *Limitations of Applicator Comparisons*

241   One difficulty and drawback of most of these studies is the patient selection. In order to gain

242   reliable and meaningful insight into dosimetric difference, both patient groups should have the same

243   tumor stage, target volume, target coverage and prescribed treatment. However, all brachytherapy

244   treatments are customized to the specific needs of the treated patient. We have also shown that HRCTV

245   and OAR dose metrics significantly vary between treating physicians (23), demonstrating that this could

246   be another confounder. The use of knowledge-based dose prediction models could overcome patient

247   selection bias by predicting the potential dose for both applicators for each patient. However, since the

248   models are trained on a certain patient cohort, this cohort can influence on the prediction accuracy of the

249   models. The effect of the different models can be seen when comparing the results of both models applied

250   to cases with the other applicator. For instance, when applying the T&R model to T&O cases, the

251   predicted T&R sparing for sigmoid was 0.15Gy per brachytherapy fraction; when applying the T&O

252   model to T&R cases, the predicted T&R sparing was 0.46Gy. Although these differences were fairly

253   small (0.16Gy (1% relative to prescription) for bladder, 0.28Gy (2%) for rectum and 0.31Gy (5%) for

254   sigmoid per brachytherapy fraction), they can be explained by a number of factors. For one, the models

255   were trained on and applied to different patient cohorts in the two scenarios, and thus could be influenced

256   by variability in clinical factors, preferences and practices of the treating physician, etc. The model-

257   predicted applicator dose differences are also confounded by model bias. For example, the T&R model

258   for sigmoid was found to have a 0.21Gy bias in the validation cohort, which means that predictions

259   tended to be 0.21Gy higher, on average, than actual values. In contrast, the T&O sigmoid model featured

260   no bias. Thus, if this bias was removed from the T&R predictions on the T&O cohort, the model-

11

261   predicted T&R sparing would increase to 0.36Gy, which is closer to the 0.46Gy sparing predicted by the

262   T&O model applied to T&R cases.

263   The way that dose is reported can also influence comparisons between applicator groups. We

264   reported dose differences in both absolute and relative dose per brachytherapy fraction, as well as total

265   EQD2 dose for completeness and comparison to prior studies, which used different dose quantities. While

266   EQD2 is the metric most commonly used to evaluate brachytherapy plan quality and accounts for

267   differences in brachytherapy prescription, results could be confounded by potential differences in EBRT

268   dose between patient groups (though EBRT dose was similar between our cohorts). Absolute

269   brachytherapy dose better highlights the difference between brachytherapy treatments, but results may be

270   confounded by differences in brachytherapy prescription. For instance, we found that bladder dose was

271   significantly higher for T&R in absolute dose per brachytherapy fraction (0.48Gy), and yet when reported

272   in relative brachytherapy dose or EQD2 it was insignificantly lower for T&R. This was likely due to the

273   higher median dose per fraction in T&O vs. T&R patients. Relative brachytherapy dose arguably gets

274   around both of these issues, though may potentially be less meaningful to practitioners used to evaluating

275   dose in Gy.

276

277   *Clinical Applications of Knowledge-Based Models*

278   We have demonstrated that separate models are required for T&R and T&O applicators, and this

279   is likely true of other applicators. Training separate models is somewhat time-consuming and requires a

280   sufficient sample size for each, but then dose predictions are more accurate for the implanted applicator,

281   and could be used to guide optimal treatment planning. One limitation of this study is that the models are

282   applied to anatomy as observed in imaging with the current applicator in place. Therefore, they do not

283   reflect any modifications to anatomy that would occur when another applicator is inserted, which could

284   additionally impact dose (e.g., the rectal retractor of the T&R applicator may push the rectum further

285   away than what is observed in a T&O scan, or the larger amount of vaginal packing with T&O may result

12

286 in different positioning of surrounding anatomy). This limits the ability to anticipate exactly what dose

287 would be received by an alternative applicator, but nonetheless studying the dose differences is important

288 to gain knowledge and create awareness of possible differences between applicators. In addition, model-

289 predicted applicator dose differences per brachytherapy fraction were very similar to those observed in

290 clinical plans, which provides some confidence that the models may provide a reasonable estimate of the

291 dose that could be achieved with an alternative applicator. The utility of models for applicator decision-

292 making will be explored in future work. This methodology could easily be applied to produce models for

293 other applicators and gain insight into specific dosimetric advantages or disadvantages in the future,

294 ensuring physicians can make informed applicator choices for patients.

295 **Conclusion**

296 Accurate knowledge-based dose prediction models were produced for T&R and T&O applicators

297 and applied to examine dose differences between two applicators that are often used interchangeably for

298 brachytherapy of cervical cancer. Both models and clinical treatment plan data indicated that significant

299 OAR sparing can be achieved with T&R over T&O, particularly for the rectum, despite similar or even

300 greater HRCTV coverage with the T&R applicator. While there are other clinical factors that may lead a

301 physician to selecting one applicator over the other, this data can help physicians to make more informed

302 decisions when determining the optimal applicator for a patient. Further, knowledge-based models could

303 be beneficial for standardizing and improving the quality of brachytherapy plans by providing patient-

304 specific quality control and dosimetric targets.

305

306 **Disclosure**

324 **References**

325 [1] Han K, Milosevic M, Fyles A, et al. Trends in the Utilization of Brachytherapy in Cervical Cancer
326   in the United States. *Int J Radiat Oncol* 2013; 87: 111–119.

327 [2] Mayadev J, Viswanathan A, Liu Y, et al. American Brachytherapy Task Group Report: A pooled
328   analysis of clinical outcomes for high-dose-rate brachytherapy for cervical cancer. *Brachytherapy*
329   2017; 16: 22–43.

330 [3] Ding L, Bi Z, Pan Z, et al. Brachytherapy-based radiotherapy is associated with improved survival
331   for newly diagnosed metastatic cervical cancer. *Brachytherapy* 2021; 20: 361–367.

332 [4] Tanderup K, Eifel PJ, Yashar CM, et al. Curative Radiation Therapy for Locally Advanced
333   Cervical Cancer: Brachytherapy Is NOT Optional. *Int J Radiat Oncol Biol Phys* 2014; 88: 537–
334   539.

335 [5] Biltekin F, Gultekin M, Yilmaz MT, et al. Dosimetric comparison of two different applicators and
336   rectal retraction methods used in inverse optimization-based intracavitary brachytherapy for
337   cervical cancer. *J Contemp Brachytherapy* 2020; 12: 35–43.

338 [6] Serban M, Kirisits C, de Leeuw A, et al. Ring Versus Ovoids and Intracavitary Versus
339   Intracavitary-Interstitial Applicators in Cervical Cancer Brachytherapy: Results From the
340   EMBRACE I Study. *Int J Radiat Oncol Biol Phys* 2020; 106: 1052–1062.

341 [7] Ma JK, Mourad WF, Allbright R, et al. Short-term clinical outcome and dosimetric comparison of
342   tandem and ring versus tandem and ovoids intracavitary applicators. *J Contemp Brachytherapy*
343   2015; 7: 218–223.

344 [8] Nv V, Dharmarajan A, PM S, et al. Dosimetric comparison of tandem and ovoids with tandem and
345   ring for intracavitary brachytherapy for carcinoma cervix. *Onkol Radioter* 2019; 13: 39–42.

346 [9] Levin D, Menhel J, Rabin T, et al. Dosimetric Comparison of Tandem and Ovoids vs. Tandem and
347   Ring for Intracavitary Gynecologic Applications. *Med Dosim* 2008; 33: 315–320.

348 [10] Gursel SB, Serarslan A, Meydan AD, et al. A comparison of tandem ring and tandem ovoid
349   treatment as a curative brachytherapy component for cervical cancer. *J Contemp Brachytherapy*
350   2020; 12: 111–117.

351 [11] Pötter R, Haie-Meder C, Limbergen EV, et al. Recommendations from gynaecological (GYN)
352   GEC ESTRO working group (II): Concepts and terms in 3D image-based treatment planning in
353   cervix cancer brachytherapy—3D dose volume parameters and aspects of 3D image-based
354   anatomy, radiation physics, radiobiology. *Radiother Oncol* 2006; 78: 67–77.

355 [12] Pötter R, Tanderup K, Kirisits C, et al. The EMBRACE II study: The outcome and prospect of two
356   decades of evolution within the GEC-ESTRO GYN working group and the EMBRACE studies.
357   *Clin Transl Radiat Oncol* 2018; 9: 48–60.

358 [13] Appenzoller LM, Michalski JM, Thorstad WL, et al. Predicting dose-volume histograms for
359   organs-at-risk in IMRT planning. *Med Phys* 2012; 39: 7446–7461.

360 [14] Ge Y, Wu QJ. Knowledge-based planning for intensity-modulated radiation therapy: A review of
361   data-driven approaches. *Med Phys* 2019; 46: 2760–2775.

362 [15] Shiraishi S, Moore KL. Knowledge-based prediction of three-dimensional dose distributions for
363   external beam radiotherapy. *Med Phys* 2016; 43: 378–387.

364 [16] Moore KL, Brame RS, Low DA, et al. Experience-Based Quality Control of Clinical Intensity-
365   Modulated Radiotherapy Planning. *Int J Radiat Oncol* 2011; 81: 545–551.

366 [17] Cornell M, Kaderka R, Hild SJ, et al. Noninferiority Study of Automated Knowledge-Based
367   Planning Versus Human-Driven Optimization Across Multiple Disease Sites. *Int J Radiat Oncol*
368   *Biol Phys* 2020; 106: 430–439.

369 [18] Giaddui T, Geng H, Chen Q, et al. Offline Quality Assurance for Intensity Modulated Radiation
370   Therapy Treatment Plans for NRG-HN001 Head and Neck Clinical Trial Using Knowledge-Based
371   Planning. *Adv Radiat Oncol*. Epub ahead of print 22 May 2020. DOI: 10.1016/j.adro.2020.05.005.

372 [19] Tol JP, Dahele M, Gregoire V, et al. Analysis of EORTC-1219-DAHANCA-29 trial plans
373   demonstrates the potential of knowledge-based planning to provide patient-specific treatment plan
374   quality assurance. *Radiother Oncol* 2019; 130: 75–81.

375 [20] Younge KC, Marsh RB, Owen D, et al. Improving Quality and Consistency in NRG Oncology
376   Radiation Therapy Oncology Group 0631 for Spine Radiosurgery via Knowledge-Based Planning.
377   *Int J Radiat Oncol Biol Phys* 2018; 100: 1067–1074.

378 [21] Nicolae A, Morton G, Chung H, et al. Evaluation of a Machine-Learning Algorithm for Treatment
379   Planning in Prostate Low-Dose-Rate Brachytherapy. *Int J Radiat Oncol* 2017; 97: 822–829.

380 [22] Shen C, Gonzalez Y, Klages P, et al. Intelligent inverse treatment planning via deep reinforcement
381   learning, a proof-of-principle study in high dose-rate brachytherapy for cervical cancer. *Phys Med*
382   *Biol* 2019; 64: 115013.

383 [23] Yusufaly TI, Kallis K, Simon A, et al. A knowledge-based organ dose prediction tool for
384   brachytherapy treatment planning of patients with cervical cancer. *Brachytherapy*. Epub ahead of
385   print 6 June 2020. DOI: 10.1016/j.brachy.2020.04.008.

386 [24] Covele BM, Puri KS, Kallis K, et al. ORBIT-RT: A Real-Time, Open Platform for Knowledge-
387   Based Quality Control of Radiotherapy Treatment Planning. *JCO Clin Cancer Inform* 2021; 134–
388   142.

389 [25] Jung H, Shen C, Gonzalez Y, et al. Deep-learning assisted automatic digitization of interstitial
390   needles in 3D CT image based high dose-rate brachytherapy of gynecological cancer. *Phys Med*
391   *Biol* 2019; 64: 215003.

392 [26] Damato AL, Viswanathan AN, Cormack RA. Validation of mathematical models for the
393   prediction of organs-at-risk dosimetric metrics in high-dose-rate gynecologic interstitial
394   brachytherapy. *Med Phys* 2013; 40: 101711.

395 [27] Damato AL, Viswanathan AN, Don SM, et al. A system to use electromagnetic tracking for the
396   quality assurance of brachytherapy catheter digitization. *Med Phys* 2014; 41: 101702.

397 [28] DeWitt KD, Hsu ICJ, Speight J, et al. 3D inverse treatment planning for the tandem and ovoid
398   applicator in cervical cancer. *Int J Radiat Oncol Biol Phys* 2005; 63: 1270–1274.

399 [29] Lahanas M, Baltas D, Zamboglou N. A hybrid evolutionary algorithm for multi-objective
400   anatomy-based dose optimization in high-dose-rate brachytherapy. *Phys Med Biol* 2003; 48: 399–
401   415.

402 [30] Oud M, Kolkman-Deurloo I-K, Mens J-W, et al. Fast and fully-automated multi-criterial treatment
403   planning for adaptive HDR brachytherapy for locally advanced cervical cancer. *Radiother Oncol*
404   2020; 148: 143–150.

405 [31] Lessard E, Pouliot J. Inverse planning anatomy-based dose optimization for HDR-brachytherapy
406   of the prostate using fast simulated annealing algorithm and dedicated objective function. *Med*
407   *Phys* 2001; 28: 773–779.

408 [32] Alterovitz R, Lessard E, Pouliot J, et al. Optimization of HDR brachytherapy dose distributions
409   using linear programming with penalty costs. *Med Phys* 2006; 33: 4012–4019.

410 [33] Shiraishi S, Tan J, Olsen LA, et al. Knowledge-based prediction of plan quality metrics in
411   intracranial stereotactic radiosurgery. *Med Phys* 2015; 42: 908–917.

412 [34] Nair MTK, Cheng M-C, Barker A, et al. High dose rate (HDR) brachytherapy technique: For
413   carcinoma of uterine cervix using nucletron applicators. *Med Dosim* 1995; 20: 201–207.

414

Table 1 *Summary of patient characteristics; T&R = tandem and ring; T&O = tandem and ovoids; HRCTV = high-risk clinical target volume; D_{2cc}, $_{EQD2}$ = D_{2cc} value reported brachytherapy dose in EQD2 with α/β=3; SD = standard deviation.*

Applicator	Parameter	Specification	Value
T&R	Number of patients	Total	47
	Number of fractions	Total	113
		Training cases	75
		Validation cases	38
	Tumor stage (FIGO)	I	58
		IIA	9
		IIB	36
		III	9
		IV	1
	HRCTV Volume [cc]	Mean (Range)	19.6 (4.9 – 40.2)
	HRCTV D90 [%]	Mean ± SD	112.5 ± 16.6
	HRCTV V100 [%]	Mean ± SD	94.9 ± 10.1
	Prescribed dose per fraction (Rx) [Gy]	Median (Range)	7 (5.5 – 8.5)
	D_{2cc} Bladder [Gy] (% Rx)	Mean ± SD	4.97 (70) ± 1.15 (15)
	D_{2cc} Rectum [Gy] (% Rx)	Mean ± SD	2.56 (36) ± 0.86 (12)
	D_{2cc} Sigmoid [Gy] (% Rx)	Mean ± SD	3.44 (48) ± 1.25 (18)
	D_{2cc}, $_{EQD2}$ Bladder [Gy]	Mean $_{EBERT +Brachy}$ (Mean $_{Brachy}$) ± SD	73.81 (30.35) ± 11.75
	D_{2cc}, $_{EQD2}$ Rectum [Gy]	Mean $_{EBERT +Brachy}$ (Mean $_{Brachy}$) ± SD	54.58 (11.12) ± 7.47
	D_{2cc}, $_{EQD2}$ Sigmoid [Gy]	Mean $_{EBERT +Brachy}$ (Mean $_{Brachy}$) ± SD	61.24 (17.78) ± 10.59
T&O	Number of patients	Total	36
	Number of fractions	Total	109
		Training cases	77
		Validation cases	32
	Tumor stage (FIGO)	I	41
		IIA	7
		IIB	57
		III	4
		IV	0
	HRCTV Volume [cc]	Mean (Range)	19.7 (7.7 – 65.7)
	HRCTV D90 [%]	Mean ± SD	105.6 ± 8.4

|		
HRCTV V100 [%]	Mean ± SD	93.5 ± 5.0
Prescribed dose per fraction (Rx) [Gy]	Median (Range)	6 (5.5 – 8)
D_{2cc} Bladder [Gy] (% Rx)	Mean ± SD	4.49 (71) ± 1.02 (14)
D_{2cc} Rectum [Gy] (% Rx)	Mean ± SD	3.27 (51) ± 0.91 (13)
D_{2cc} Sigmoid [Gy] (% Rx)	Mean ± SD	3.76 (60) ± 0.86 (13)
$D_{2cc, EQD2}$ Bladder [Gy]	Mean $_{EBERT+Brachy}$ (Mean $_{Brachy}$) ± SD	74.42 (30.95) ± 9.23
$D_{2cc, EQD2}$ Rectum [Gy]	Mean $_{EBERT+Brachy}$ (Mean $_{Brachy}$) ± SD	62.54 (19.07) ± 7.14
$D_{2cc, EQD2}$ Sigmoid [Gy]	Mean $_{EBERT+Brachy}$ (Mean $_{Brachy}$) ± SD	67.10 (23.63) ± 7.62

417

418

419 **Table 2** *Summary of dose differences between patients treated with T&O and T&R applicators (validation and trainings cases). μ*
420 *= average; $D_{2cc, EQD2} = D_{2cc}$ value represented in EQD2 with $α/β = 3$; **p < 0.01; *p < 0.05; (% Rx) = dose relative to prescribed*
421 *dose; [1] T&R model applied to T&O cases; [2] T&O model applied to T&R cases*

| Parameter	Clinical Differences	Predicted Differences[1]	Predicted Differences[2]
	$μ_{T\&O, actual} - μ_{T\&R, actual}$	$μ_{T\&O, actual} - μ_{T\&R, predicted}$	$μ_{T\&O, predicted} - μ_{T\&R, actual}$
D_{2cc} Bladder [Gy] (% Rx)	-0.48** (1%)	+0.33* (6%**)	+0.49** (7%**)
D_{2cc} Rectum [Gy] (% Rx)	+0.71** (16%**)	+0.79** (13%**)	+1.07** (15%**)
D_{2cc} Sigmoid [Gy] (% Rx)	+0.32* (11%**)	+0.15 (2%)	+0.46** (7%**)
$D_{2cc, EQD2}$ Bladder [Gy]	+0.61	+2.62*	+5.69**
$D_{2cc, EQD2}$ Rectum [Gy]	+7.96**	+6.15**	+7.31**
$D_{2cc, EQD2}$ Sigmoid [Gy]	+5.86**	+0.69	+3.65**

422

423
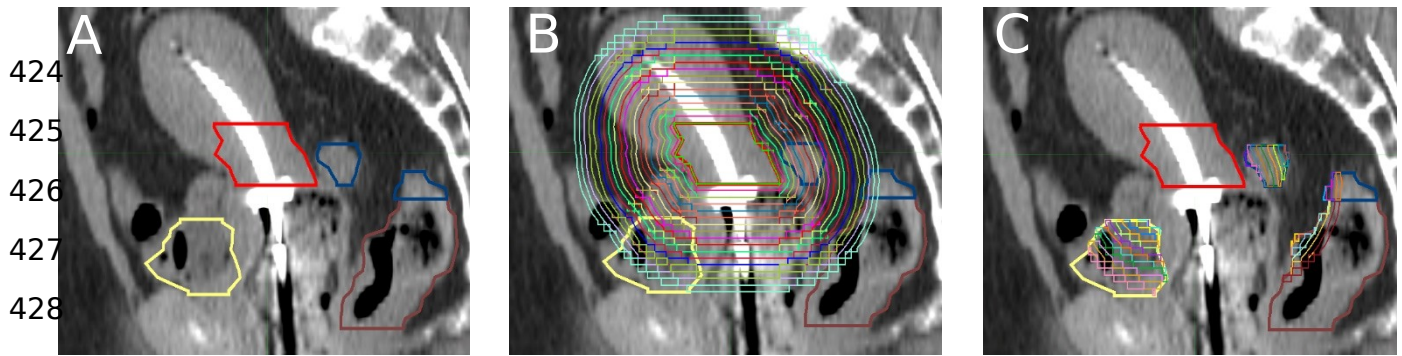
Figure 1 Pre-processing required for knowledge-based dose predictions. First, target and OARs are contoured (example T&O CT sagittal slice, A). Then shells are generated around the target and represent distance from target (B). Finally, dose is extracted from each shell where it overlaps with each OAR (C), and used to generate DVH dose prediction models. Steps B onwards were fully automated within MIM.

433
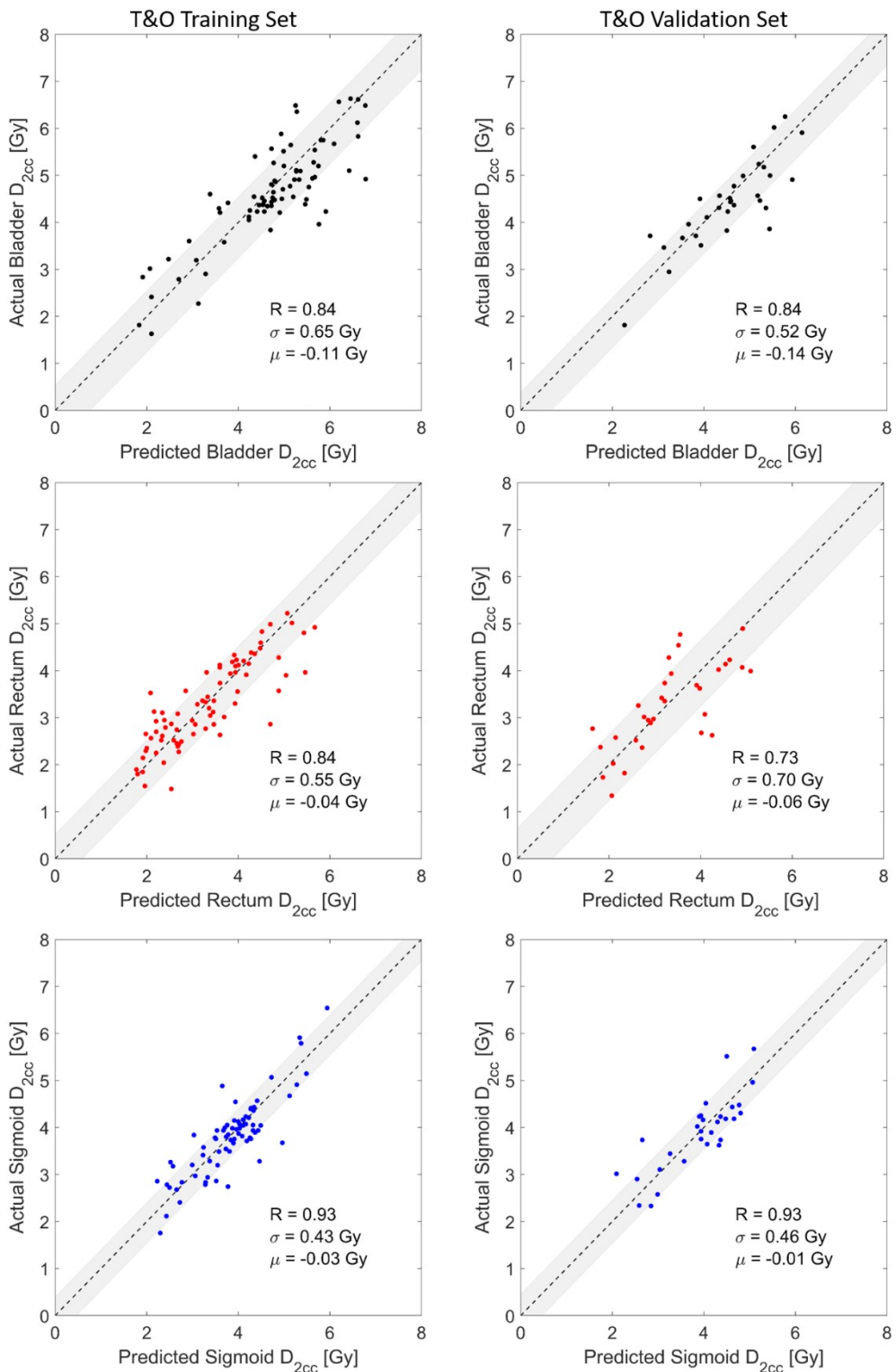
Figure 2 *Actual verses predicted D_{2cc} values for each OAR for training (left) and validation (right) data sets for the T&O model. Pearson correlation coefficients (R), standard deviation (indicated by σ as well as gray color wash, representing model precision), and mean (μ, representing model bias) of $\Delta D_{2cc} = D_{2cc, actual} - D_{2cc, predicted}$ are shown. Black lines indicate hypothetical perfect model predictions.*

438
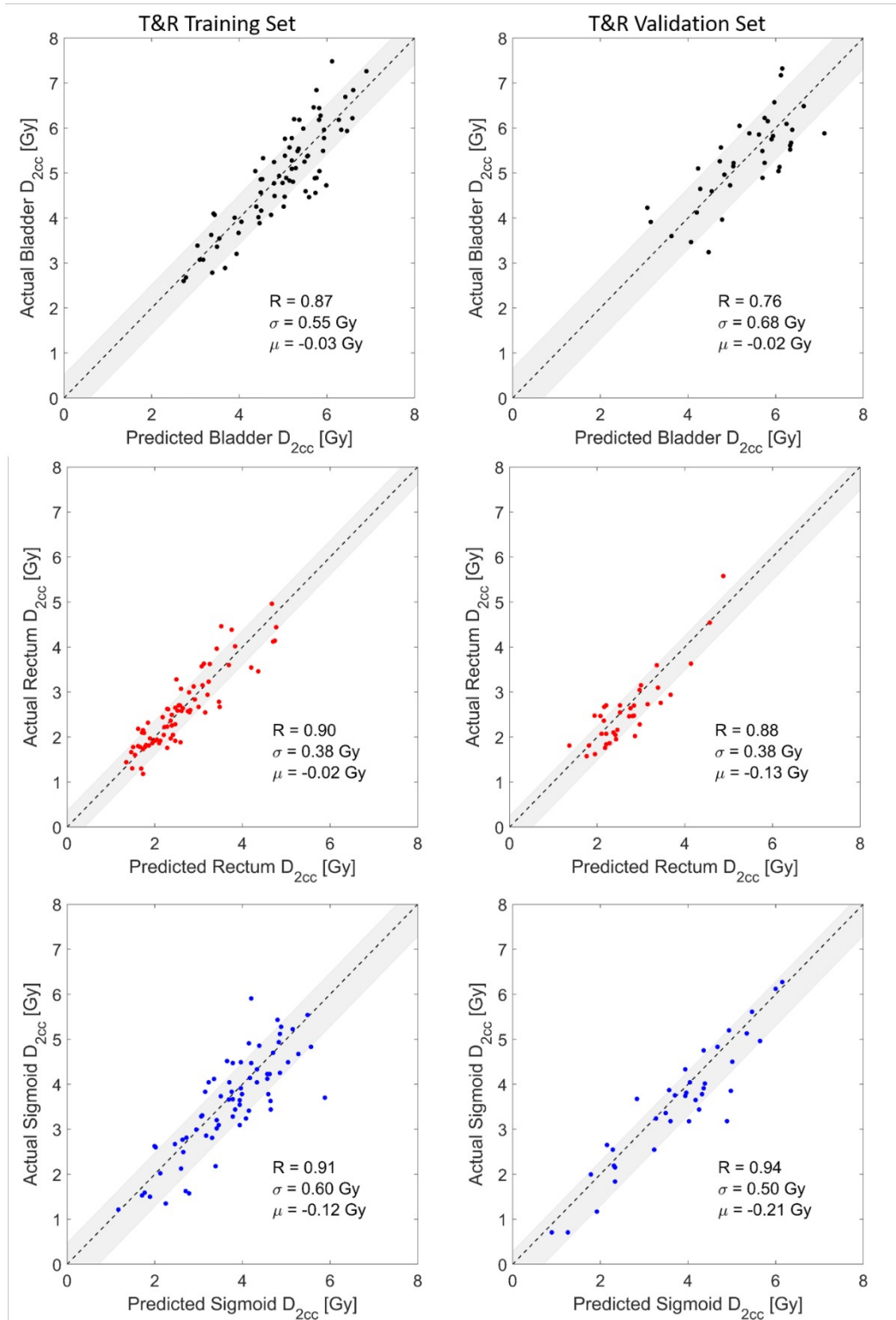
Figure 3 *Actual verses predicted D_{2cc} values for each OAR for training (left) and validation (right) data sets for the T&R model. Pearson correlation coefficients (R), standard deviation (indicated by σ as well as gray color wash, representing model precision), and mean (μ, representing model bias) of $\Delta D_{2cc} = D_{2cc,\ actual} - D_{2cc,\ predicted}$ are shown. Black lines indicate hypothetical perfect model predictions.*

443
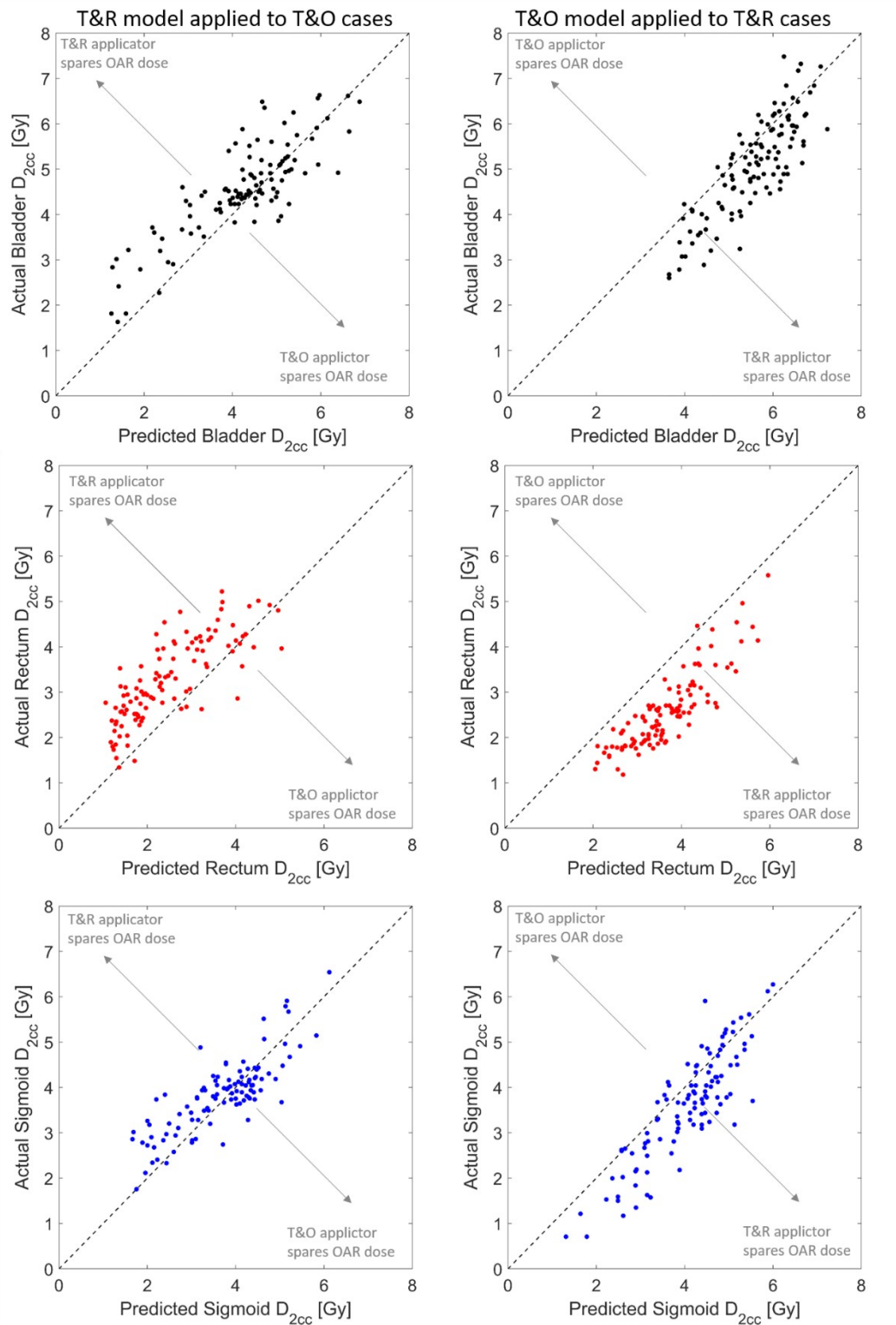
Figure 4 Actual dose for T&O cases vs. D_{2cc} values predicted by T&R model (left), and actual dose for T&R cases vs. T&O model predictions (right). All patients from training and validation datasets were included. Black dotted line indicates equivalence between model predictions and actual values. A bias towards one side of the line marks possible dose differences between the applicators.