

1 **Purpose:** The use of interstitial needles, combined with intracavitary applicators, enables customized dose
2 distributions and is beneficial for complex cases, but increases procedure time. Overall, applicator selection is not
3 standardized and depends on physician expertise and preference. The purpose of this study is to determine whether
4 dose prediction models can guide needle supplementation decision-making for cervical cancer.

5 **Materials and Methods:** Intracavitary knowledge-based models for organ-at-risk (OAR) dose estimation were
6 trained and validated for tandem-and-ring/ovoids (T&R/T&O) implants. Models were applied to hybrid cases with
7 1-3 implanted needles to predict OAR dose without needles. As a reference, 70/67 hybrid T&R/T&O cases were
8 replanned without needles, following a standardized procedure guided by dose predictions. If a replanned dose
9 exceeded the dose objective, the case was categorized as requiring needles. Receiver operating characteristic (ROC)
10 curves of needle classification accuracy were generated. Optimal classification thresholds were determined from the
11 Youden Index.

12
13 **Results:** Needle supplementation reduced dose to OARs. However, 67%/39% of replans for T&R/T&O met all dose
14 constraints without needles. The ROC for T&R/T&O models had an area-under-curve of 0.89/0.86, proving high
15 classification accuracy. The optimal threshold of 99%/101% of the dose limit for T&R/T&O resulted in
16 classification sensitivity and specificity of 78%/86% and 85%/78%.

17 **Conclusion:** Needle supplementation reduced OAR dose for most cases but was not always required to meet
18 standard dose objectives, particularly for T&R cases. Our knowledge-based dose prediction model accurately
19 identified cases that could have met constraints without needle supplementation, suggesting that such models may be
20 beneficial for applicator selection.

21

22 **Keywords:** knowledge-based; dose prediction; cervix cancer, brachytherapy; needle supplementation

23

24

25

26

27

28

29

30

31 **Introduction**

32 A brachytherapy boost after external beam radiation (EBRT) is considered to be the standard of care for
33 cervical cancer patients (1, 2). Brachytherapy is often delivered using intracavitary applicators such as tandem-and-
34 ring (T&R) or tandem-and-ovoids (T&O) (3). For more complex cases, for example those with large or asymmetric
35 high-risk clinical target volumes (HRCTVs), interstitial needles are often added to better customize the dose
36 distribution, reducing dose to organs-at-risk (OARs), and/or improving HRCTV coverage (4–6). However, using
37 needles increases procedure time and complexity (7). Additionally, the decision to use needles is not standardized
38 and depends on the expertise and preference of the physician. The increased procedural complexity and need for
39 expertise may lead to barriers to brachytherapy implementation, which is problematic given that brachytherapy
40 utilization is diminishing despite its proven benefit (2).

41 Knowledge-based models for plan quality analysis or automated treatment planning have been extensively
42 studied in EBRT (8–10), with over a decade of development, many publications and existing commercial solutions.
43 However, exploration of these methods are comparably rare in brachytherapy (11–13). Knowledge-based models
44 use prior information from treatment plans to predict dose for new cases based on contours, distances from the
45 target-to-OARs, and other geometric features (9).

46 Various EBRT studies have proven that the use of knowledge-based models improves plan quality, reduces
47 planning variability and leads to more standardized plans with greater OAR sparing (14–17). Similar to EBRT,
48 knowledge-based models could help to standardize future brachytherapy treatment planning and predicted doses
49 could serve as quality assurance metrics. However, unlike EBRT, brachytherapy plan quality is largely dictated by
50 the applicators and/or needles implanted. There are a variety of different applicators to choose from and strategies
51 for implantation, which leads to large plan quality variations when coupled with variability in treatment planning.
52 Knowledge-based models could play a unique role in brachytherapy, where dose predictions for different treatment
53 types could help to inform the optimal treatment technique.

54 The purpose of this study is to determine whether knowledge-based models can predict cases where needle
55 supplementation is required to meet dose objectives. To investigate this, existing intracavitary models were applied
56 to hybrid cases treated with additional needles, and estimated OAR doses were used to predict cases that required
57 needle supplementation. Accuracy of classification was verified by replanning all cases without needles, and

58 compared to the current classification in our clinic. Knowledge-based dose estimations could form the basis for
59 more objective, standardized decision-making on necessity of needle usage.

60 **Materials and Methods**

61 *Patient Cohort*

62 Cervical cancer patients treated between 2017-2020 with image-guided high-dose-rate brachytherapy with
63 T&O or T&R applicators and 0-3 implanted interstitial needles were included (UCSD IRB #200065C). Treatment
64 planning, applicator reconstruction and contouring were performed based on computed tomography (CT) imaging.
65 Cases implanted with both an intracavitary applicator and needles are referred to as “hybrid” cases. A “case” in this
66 context is defined as a single fraction of a brachytherapy treatment, consisting of 3-5 fractions. Detailed information
67 about the patient cohort is presented in Table 1. Organs and HRCTV (cervix, gross disease and any potential
68 extension) were contoured on the CT. Clinical treatment planning was conducted according to most of the EQD2
69 dose objectives defined by EMBRACE II (18, 19), which include limits (HRCTV D90>85Gy, bladder D_{2cc}<90Gy,
70 rectum D_{2cc}<75Gy and D_{2cc} sigmoid<75Gy) and soft planning aims (recommended but not required, bladder
71 D_{2cc}<80Gy, rectum D_{2cc}<65Gy, and sigmoid D_{2cc}<70Gy). After normalization to point A, plans were optimized to
72 meet objectives and minimize OAR dose. During this process, the overall EQD2 for HRCTV D90 and organ D_{2cc}
73 values were projected using a spreadsheet and then compared to the dose objectives, to ensure that the treatment
74 plan fulfilled all necessary dose criteria. All reported clinical parameters, e.g., D90, V100 and D_{2cc}, were exported
75 directly out of the treatment planning system. All patients received prior EBRT, and the most common EBRT
76 prescription was 45Gy in 25 fractions.

77 *Knowledge-Based Dose-Prediction Models*

78 Previously trained and validated dose-prediction models for intracavitary applicators were used. Full details
79 about the algorithm can be found in Yusufaly *et al.* (13), which demonstrates the accuracy of dose-prediction
80 models for cervical cancer patients treated with T&O. Briefly, DVH models are 1D radial models, which predict
81 dose to OARs based on the radial distance from an OAR sub-volume to the HRCTV. OARs are divided into sub-
82 volumes based on the boundary distance to the HRCTV contour by overlapping the OAR contour with 2mm shells
83 expanding around the HRCTV. A differential DVH for each OAR shell is extracted from each treatment plan in the
84 training set, and averaged to produce dose kernels as a function of radial distance. For model predictions, OAR

85 contours are discretized in the same manner, and the differential DVH of the considered OAR is calculated from the
86 sum of differential DVH kernels, weighted by the volume of each OAR sub-volume (8, 13, 20). DVHs for new
87 patients can be predicted using HRCTV and OAR contours alone. Additional details about models, are described in
88 the Supplementary Materials. Two different models were trained for T&R or T&O applicators, since these
89 applicators result in significantly different OAR dose (21, 22). The T&R (T&O) model was trained on 75 (80)
90 intracavitary cases and validated on 38 (32) cases (see Table 1 for more details), achieving an approximately
91 70/30% split between training/validation.

92 Since OAR D_{2cc} is a common metric used to evaluate brachytherapy plan quality, these metrics were
93 extracted from all DVHs. Unless otherwise stated, D_{2cc} values in results and figures are reported in absolute dose per
94 brachytherapy fraction in Gy. Model performance was quantified using the difference between actual D_{2cc} s obtained
95 from clinical plans and the predicted D_{2cc} s, i.e. ΔD_{2cc} , as well as the standard deviation over these $\sigma_{Model}(\Delta D_{2cc})$,
96 representing model precision. The T&R (T&O) σ_{Model} was 0.66Gy (0.52Gy), 0.39Gy (0.70Gy) and 0.50Gy (0.46Gy)
97 for bladder, rectum and sigmoid, respectively (see Supplementary Figure 4).

98 Model training, validation and extraction of predicted DVHs were performed using automated scripts in
99 MIM (v7.0.1, MIM Software Inc., Cleveland, OH). D_{2cc} calculation and data analysis was performed using
100 MATLAB (R2019b, MathWorks, Inc., Natick, MA).

101 *Replanning Procedure*

102 Seventy (67) hybrid T&R (T&O) cases were replanned without loading the inserted needles to demonstrate
103 the achievable dose using an intracavitary applicator alone (see Supplementary Figure 1, which describes the
104 standardized replanning procedure). This process mimics manual, clinical treatment plan optimization with the
105 addition of dose predictions to guide planning. Patient-specific dose objectives were calculated based on the total
106 number of prescribed fractions and dose as well as the dose received in prior EBRT, to ensure EQD2 objectives
107 were met. First, the clinical treatment plan was reset with standard loading of the intracavitary applicator. After
108 normalizing to point A, the dose distribution was shaped to cover the HRCTV. If the first mandatory aim, HRCTV
109 $D_{90} \geq 85\text{Gy}$ EQD2, was fulfilled then the plan was further tuned to reduce OAR dose, attempting to meet D_{2cc}
110 constraints. The predicted doses were used to guide optimization; however, the aim was to minimize OAR dose, i.e.,
111 when possible, the plan was further improved until D_{2cc} values were lower than the predicted D_{2cc} values $\pm \sigma_{Model}$. In

112 order to quantitatively replicate the visual evaluation that occurs while planning at our center (i.e., covering the
 113 HRCTV target with the prescription isodose) we also attempted to meet a third criteria, $V_{100} \geq 95\%$ (where 100%
 114 corresponds to the prescribed dose per brachytherapy fraction as shown in Table 1), although this was prioritized
 115 last. All replans were created by a trained postdoctoral fellow and verified and/or further modified by experienced
 116 brachytherapy physicists (DB, DS, KK, XR). All treatment planning was performed in the same brachytherapy
 117 planning system (BrachyVision, v15.6, Varian Medical Systems, Inc., Palo Alto, CA). Replanned doses were
 118 compared to predicted and actual hybrid dose values using a two-sample Student's t-test (significance level of 0.05).

119 *Receiver-Operating-Characteristic Analysis*

120 In order to quantify the ability of the models to accurately identify cases that require needles, receiver
 121 operating characteristic (ROC) curves were constructed and the area under the curve (AUC) determined. The "true
 122 condition" was defined by the replan: any intracavitary plan with an OAR dose exceeding the objective was
 123 identified as a case requiring needles. For the model, a threshold was used for needle classification. This threshold
 124 was taken as a percentage of the dose objective for an OAR, and was varied from 50% to 200% to reconstruct the
 125 ROC plot. Since all dose objectives are based on EQD2 including EBRT, we computed a patient-specific objective
 126 for a single brachytherapy fraction by assuming that all brachytherapy fractions received the same dose. Analysis
 127 was performed for two different sets of objectives – the dose limits and soft planning aims - since some clinics may
 128 use the more conservative aims to determine when to implant needles. True positives were defined as needle cases
 129 (any replanned $OAR > \text{dose objective}$) that were correctly identified as such by the model (any predicted OAR
 130 $\text{dose} > \text{threshold}$). Sensitivity and specificity of the model classification were then calculated for each OAR separately
 131 and combined, using the following equations, and positive/negative predictive value (PPV/NPV) were computed.

$$132 \text{ Sensitivity} = \frac{\text{\# cases with } D_{2cc, \text{replan}} \geq \text{dose objective} \wedge D_{2cc, \text{predict}} \geq \text{threshold}}{\text{\# cases with } D_{2cc, \text{replan}} \geq \text{dose objective}}$$

$$133 \text{ Specificity} = \frac{\text{\# cases with } D_{2cc, \text{replan}} < \text{dose objective} \wedge D_{2cc, \text{predict}} < \text{threshold}}{\text{\# cases with } D_{2cc, \text{replan}} < \text{dose objective}}$$

134 Because the hybrid dataset is biased towards cases requiring needles, an equivalent number of randomly
 135 selected intracavitary cases (70 T&R and 67 T&O) were added to the classification analysis. Any plans that did not
 136 fulfill HRCTV $D_{90} 85\text{-}90\text{Gy}$ were rescaled to ensure equivalent coverage to replans. Our current clinical status for

137 needle classification was also determined retrospectively for comparison, which is based on physician experience
138 using knowledge from prior fractions (if any), clinical examination, D90 planning goals based on the number of
139 planned fractions, or imaging. A patient that was actually treated with needles was classified by the physician as
140 needing needles, and vice versa for an intracavitary applicator alone. The optimal threshold for model classification
141 was determined using the Youden index, which maximizes sensitivity and specificity (23). The optimal threshold is
142 defined as the point on the ROC curve that corresponds to the maximal Youden index.

143 **Results**

144 *Comparison of Intracavitary Replans to Actual Clinical Hybrid Plans*

145 Organ doses for intracavitary replans are compared to values from clinical hybrid plans in Figure 1.
146 Overall, OAR dose sparing was improved with interstitial needles, especially bladder dose when using T&O. For a
147 single brachytherapy fraction, clinical hybrid implant D_{2cc} for bladder, rectum and sigmoid were, on average,
148 5.23 ± 1.01 Gy (5.18 ± 0.84 Gy), 2.81 ± 0.93 Gy (4.07 ± 0.96 Gy), and 4.00 ± 1.04 Gy (3.89 ± 1.05 Gy), for T&R (T&O)
149 applicators. The replanned OAR doses were higher, on average, with mean T&R (T&O) D_{2cc} of 6.02 ± 2.33 Gy
150 (6.53 ± 1.88 Gy), 2.95 ± 0.94 Gy (4.69 ± 1.86 Gy), and 3.91 ± 0.74 Gy (4.32 ± 1.61 Gy) for bladder, rectum and sigmoid,
151 respectively. Replanned and clinical doses were significantly different for bladder (mean deviation=1.11 Gy and
152 $p=0.01$ for T&R, 1.38 Gy and $p<0.001$ for T&O) and rectum for T&O (mean deviation=0.83 Gy, $p=0.02$).

153 Using our standardized procedure for replanning, a mean and standard deviation of the HRCTV D90 of
154 $107 \pm 6\%$ ($105 \pm 7\%$) was obtained for all T&R (T&O) cases (see Supplementary Figure 3), where dose is relative to
155 the prescription per brachytherapy fraction. The clinical plans featured greater variation in target coverage, with
156 HRCTV D90 values of $114 \pm 13\%$ ($105 \pm 10\%$) for T&R (T&O) applicators. The replans indicated that 67% (39%) of
157 the hybrid T&R (T&O) cases would have met the OAR dose limits and 43% (9%) would have met the softer
158 planning aims without needles.

159 *Comparison of Predicted and Replanned Intracavitary Dose*

160 The predicted D_{2cc} were, on average for T&R (T&O), 6.02 ± 2.03 Gy (6.68 ± 1.16 Gy), 3.00 ± 0.86 Gy
161 (4.70 ± 0.75 Gy), and 4.40 ± 0.79 Gy (4.33 ± 0.87 Gy) for bladder, rectum and sigmoid respectively. A comparison of
162 predicted and replanned doses is shown in Figure 2 and Supplementary Figure 3. As can be seen in Figure 2 and

163 Table 2, most of the sigmoid and rectum replanned doses fell within precision of the model predictions (i.e. within
164 the grey band), or below. The results for bladder were slightly more variable, particularly for cases exceeding the
165 dose limits, but most cases fell within the diagonal quadrants (indicating that both replan and predicted doses were
166 either greater or less than the dose limits). Besides the sigmoid for the T&R applicator ($p < 0.001$), the replanned D_{2cc}
167 did not significantly differ from predicted D_{2cc} values. The mean \pm standard deviation of differences between
168 replanned and predicted D_{2cc} (i.e. $D_{2cc, \text{replanned}} - D_{2cc, \text{predicted}}$) were $0.00 \pm 0.95 \text{ Gy}$ ($-0.15 \pm 1.52 \text{ Gy}$) for bladder, -
169 $0.06 \pm 0.43 \text{ Gy}$ ($-0.01 \pm 1.56 \text{ Gy}$) for rectum and $-0.49 \pm 0.47 \text{ Gy}$ ($-0.01 \pm 1.19 \text{ Gy}$) for sigmoid for T&R (T&O). As shown
170 in Table 2 and Figure 2, there were very few cases ($\leq 10\%$) for which the replanned dose exceeded the dose limit,
171 while the predicted dose met the limit (top left quadrant of plots in Figure 2). Considering all OARs together, 77%
172 (68%) of T&R (T&O) cases had replanned D_{2cc} values lower than or within the uncertainty of the predictions.

173 *Needle Classification Performance Using Dose Limits*

174 ROC curves for model needle classification are displayed in Figure 3. The AUC using dose limits as
175 decision criteria for needle supplementation were 0.89 (0.86) for T&R (T&O), indicating good model prediction
176 accuracy. These plots use all organ doses to classify needle cases, while AUC metrics for a single organ are
177 displayed in Table 2. The classification accuracy of our clinic based on the physician's choice of needle usage (red
178 crosses on Figure 4) corresponds to a sensitivity of 72% (73%) and specificity of 56% (67%) for T&R (T&O),
179 which was lower than what was theoretically achieved with the predictions as guidance (Figure 3A and B). The
180 optimal threshold was determined to be 99% of the dose limit for T&R (providing a sensitivity of 78% and
181 specificity of 85%), and 101% of the dose limit for T&O (providing a sensitivity of 86% and specificity of 78%)
182 (black crosses on Figure 4). This corresponded to a PPV and NPV of 61% (74%) and 93% (88%) for T&R (T&O).

183 *Needle Classification Performance Using Planning Aims*

184 Classification results using planning aims as the decision boundary for needle supplementation were similar. AUC
185 values were 0.88 (0.89) for T&R (T&O). The sensitivity and specificity of the clinical classification were
186 determined to be 66% (60%) and 34% (81%) for T&R (T&O), respectively (Figure 3C and D). The optimal
187 threshold for model classification was 104% of the planning aim for T&R (sensitivity=84% and specificity=81%)

188 and 110% for T&O (sensitivity=79% and specificity=91%). This corresponded to a PPV and NPV of 77% (96%)
189 and 86% (58%) for T&R (T&O) cases.

190 **Discussion**

191 This study explores the benefits of knowledge-based models to guide brachytherapy needle
192 supplementation decision-making. Although there are solutions for automating various aspects of brachytherapy
193 treatment planning, such as applicator reconstruction (24–26) and inverse optimization (12, 27), there currently are
194 no tools to guide gynecological applicator choice. Therefore, the addition of interstitial needles is not standardized
195 and relies on the physician’s preference and expertise, which is particularly challenging for inexperienced
196 physicians. The proposed model predictions could help physicians make informed decisions based on quantitative
197 metrics.

198 *Relative Benefit of Needles vs. Standard Applicators*

199 In this series, we found that needles were not always needed to meet the dose objectives. 67% (39%) of the
200 T&R (T&O) hybrid cases could have met the OAR dose limits with an intracavitary applicator alone and 43% (9%)
201 of the T&R (T&O) hybrid cases would have even met the planning aims. However, the use of needles can enable
202 further OAR sparing and/or target dose escalation (28), and sometimes physicians will add needles to achieve either
203 of these goals based on patient and tumor anatomy. The overall objective of treatment planning is to get the best
204 coverage with the lowest OAR doses, and physicians in our clinic often add needles to meet softer planning aims
205 rather than dose limits. Our data showed that supplemental needles resulted in significant dose reductions for
206 bladder (0.78-1.35Gy on average) and rectum (0.43Gy, T&O only). The bladder seemed to benefit the most from
207 needles; for 65% (34%) of T&R (T&O) cases, bladder was the only organ that exceeded the dose limit, indicating
208 that the bladder was the limiting factor for needle requirement. Overall, the hybrid T&O cases had a larger HRCTV
209 volume in comparison to the hybrid T&R cases (see Table 1), which might be linked to the greater percentage of
210 T&O cases requiring needles. Hybrid plans featured a wider range of target coverage than our standardized
211 intracavitary replans, though this may reflect different patient-specific clinical requirements, as opposed to non-
212 standardization of planning or applicator differences. The slightly small HRCTV volumes in the hybrid patient

213 group (see Table 1) are likely due to the fact that purely interstitial cases and hybrid cases with more than 3 needles
214 were excluded.

215 It is undeniable that needle supplementation can improve dosimetry, but it comes at the cost of increased
216 procedure time and complexity, which is potentially a barrier to institutions that lack brachytherapy expertise or
217 resources. The analysis and tools presented in this study could allow physicians to weigh the relative benefit of
218 needles vs. standard applicators, and to make more informed decisions about needle supplementation that are
219 appropriate for their center.

220 *Model Dose-Prediction Accuracy*

221 Dose-prediction models were fully implemented with scripting in MIM, meaning that once images were
222 contoured, OAR dose predictions could be computed within seconds. The use of CT-based planning as opposed to
223 magnetic resonance (MR)-based should not impact results, since the models rely only on the target and organ
224 contours and were trained on cases with a wide range of HRCTV volumes. However, future work will examine
225 model accuracy at institutions that use MR for contouring.

226 This manuscript mainly focuses on the application of knowledge-based models to needle discrimination,
227 but the predicted doses could also be beneficial for patient-specific plan quality control during treatment planning.
228 Model prediction precision ranged from 0.39-0.70Gy for intracavitary validation cases. Although the models were
229 applied to more complex cases, the model performance was still reasonable. Replanned D_{2cc} values were within the
230 precision of T&R (T&O) model-predictions or lower for 77% (68%) of cases considering all OARs. This result is
231 surprising, given that the model was trained on intracavitary cases, where organ doses rarely exceeded the limits and
232 the anatomy was less challenging. We suspect that prediction accuracy would improve if we added more complex
233 cases to the training set, which would now be possible since we have intracavitary replans for all these patients. The
234 fact that many replanned OAR doses were lower than predicted $D_{2cc} - \sigma_{Model}$ (see Figure 2) is expected, because the
235 models were produced using variable clinical plans, while replanning was performed using a standardized workflow
236 guided by predictions. This effect has been observed in EBRT studies as well, where using knowledge-based
237 planning significantly reduced OAR dose (29). The current model predictions are most appropriate for quality
238 control of intracavitary treatment plans. Nonetheless, future work will develop separate models for hybrid and

239 interstitial cases in order to provide more accurate D_{2cc} predictions and valuable quality control for more complex
240 cases.

241 *Needle Classification Accuracy*

242 The ROC analysis and AUCs, ranging between 0.88-0.89 (T&R) and 0.86-0.89 (T&O), demonstrated that
243 knowledge-based models can effectively discriminate between cases that require needles and cases for which an
244 intracavitary applicator is sufficient, with high sensitivity and specificity. The predicted classification was beneficial
245 for needle supplementation guidance when reviewed retrospectively and not considering patient-specific needs.
246 Using the planning aims to guide needle decision-making, which more closely reflects our clinical practice, the
247 current sensitivity of needle selection is 66% (60%) and specificity is 62% (81%) for T&R (T&O), but includes
248 other clinical factors not represented in the model. Dose predictions could improve sensitivity and specificity for
249 T&R (T&O) to 84% (79%) and 81% (91%), respectively, using the optimal threshold of 104% (110%) of planning
250 aims to indicate needle requirement. In practice, this means that one could run the model for a given T&R case and
251 if predicted doses exceeded 104% of the planning aims, needles would be recommended. There would be an 84%
252 chance that a case requiring needles would be accurately identified as such by the model, and an 81% chance that a
253 case would be accurately identified as not requiring needles, if other clinical determinants did not favor needle use.
254 We suspect that the current model predictions (both doses and classification) would be meaningful for any
255 institution using standard T&R or T&O applicators and similar treatment planning criteria. Future work will test the
256 accuracy and applicability of these models on datasets and workflows of independent institutions.

257 *Standardized Planning*

258 The standardized planning workflow, driven by knowledge-based dose estimations, was another useful
259 outcome from this work, and could be beneficial to implement in clinical practice. Using the predictions and dose
260 criteria as guidance, a postdoctoral fellow was easily trained to produce clinically acceptable plans in reasonable
261 timeframes (about 10 minutes, not taking catheter digitization and quality assurance into account). Only about 10%
262 of the cases required adjusting after review by experienced brachytherapy physicists. This suggests that guiding
263 treatment planning with predictions could be beneficial for training residents and inexperienced clinicians, and could
264 potentially speed up treatment planning for experienced clinicians. Using patient-specific dose predictions as

265 targets, the planner could push harder on organ dose than they otherwise might have under time constraints, which is
266 usually one limiting factor of plan optimization. This same effect has been reported for EBRT: when only static
267 population-based limits were used to guide planning, high degrees of plan quality variability and excess dose to
268 normal tissues were observed, while when patient-specific achievable dose limits were provided, variability and
269 normal tissue doses were substantially reduced (30, 31). Standardized planning, guided by dose predictions, could
270 help physicians and physicists create optimal treatment plans for patients and standardize plan quality.

271 *Proposed Clinical Workflow Using Needle Supplementation Guidance*

272 Models are most accurate for same-day predictions, and thus would be very beneficial for centers with
273 image-guided brachytherapy suites (see Figure 4A for proposed workflow in clinical implementation). In this case,
274 an intracavitary applicator should be implanted for each fraction, and after imaging and contouring, the models
275 should be used to predict if needles should be added. Adding needles and re-imaging after initial imaging may only
276 be feasible at a minority of centers (32), so we also propose the following alternative workflow (see Figure 4B). For
277 the first fraction, prior knowledge from examination and imaging should be used to decide if needle supplementation
278 is necessary, prioritizing the use of purely intracavitary applicators. If the patient anatomy includes specific features,
279 such as irregular tumor topography, tumor size >4cm, bulky parametria disease or posterior bulky disease, needles
280 should be implanted. After contouring on imaging, model-based predicted organ doses should be obtained. If any
281 predicted organ dose exceeds 99% (101%) of the dose limit for T&R (T&O), the next fraction should be
282 supplemented with needles. Physicians that prefer to implant needles to meet the planning aims would instead
283 implant needles if any predicted OAR dose exceeded 104% (T&R) or 110% (T&O) of the planning aim. Although
284 actual organ doses are available after treatment plan optimization, we still recommend using model predictions to
285 guide decision-making in case treatment planning was suboptimal. This procedure should be repeated for each of the
286 following fractions, such that the applicator decision is guided by dose predictions of the previous fraction.
287 Physicians who are more comfortable with needle usage and prefer to err on the side of caution could choose to
288 implant needles more often for the first fraction; however, the following workflow determining needle requirement
289 for the following fraction could still apply. Finally, for any fraction treated without needles, we would recommend
290 aiming for the predicted organ doses (minus model precision) when treatment planning, and comparing the final
291 organ doses to predicted values as a means of quality control.

292 Using the models provides the following benefits to clinicians over current clinical practice. For one,
293 current treatment planning is performed under time pressure, and thus clinicians may not be able to easily determine
294 the optimal dose for a patient in the allotted time. For purely intracavitary implants, a clinician may believe that
295 needles are needed for the next fraction, while more optimal treatment planning could have met constraints. The
296 models provide reasonable estimates of achievable OAR dose with intracavitary applicators, which could speed up
297 the planning process, provide additional assurance that the treatment plan is optimal and confirmation as to whether
298 needles will be needed in future fractions. A physician may choose to insert needles for one fraction, but without
299 replanning the case without needles, they may not know for sure whether needles were actually necessary to meet
300 dose constraints. Our own data shows that needles were over-used, and thus there is room for improvement even
301 within an experienced, busy brachytherapy clinic. The models can provide additional guidance on what to do for the
302 next fraction on these borderline cases.

303 *Limitations and Future Work*

304 The models were trained on clinical treatment plans, but brachytherapy plans are not homogenous and a
305 decent amount of variation between plans, individualized for the needs of each patient, is to be expected. This means
306 some variability is incorporated into the models; however, we expect that much of this variation averaged out over
307 the large patient cohort. The models are simple and make assumptions about dose conformality to the HRCTV,
308 which do not perfectly hold for the standard pear-shaped distributions of cervical brachytherapy dose. The models
309 also are less accurate when the tandem is not centered on the HRCTV, which necessitated an additional correction
310 for HRCTV asymmetry to improve bladder dose predictions (see Supplementary Materials). Despite these
311 limitations, the models were accurate for needle discrimination in their current form. Future work will explore 3D
312 dose prediction, which can account for greater complexities in anatomy and applicator geometry.

313 One limiting factor is that dose-prediction models assume standard situations and do not account for any
314 dose received in previous fractions or other individual patient needs (such as additional dose to the vagina). We
315 recognize physicians consider additional clinical data and the full patient history during treatment planning. For
316 instance, physicians may opt to increase target coverage, and this is not reflected in either the replanning or
317 predicted doses. These factors influence of the clinical classification and retrospective analysis of sensitivity and
318 specificity. To provide a more fair comparison, a prospective evaluation of physician classification based solely on

319 whether OARs are anticipated to meet dose objectives (i.e. without HRCTV dose escalation) is needed. This will be
320 explored in future work. However, the models indicate what kind of dose is achievable if target coverage of D90 85-
321 90Gy is met, and whether needles are required to meet dose objectives. Physicians could use this information to
322 guide decision-making for a given case alongside other characteristics they value. If the models are employed at the
323 first fraction, inter-fraction variability may be reduced, and the need to compensate for non-ideal dose from prior
324 fractions could be alleviated.

325 The addition of the HRCTV V100 criteria to replanning objectives meant that HRCTV D90 could take on a
326 range from 85 to 90 Gy EQD2. As a result, patients with different prescriptions could have slightly different target
327 coverage, which could impact both organ D_{2cc} and needle classification in replans. In order to assess whether
328 prescription was influencing results, we compared the following parameters between patients treated with different
329 fractionation schemes using a Kruskal-Wallis test with a Bonferroni correction: $D_{2cc, replanned} - \text{dose limit}$ and
330 $D_{2cc, replanned} - D_{2cc, predicted}$. A chi-squared test of independence was used to determine whether fractionation scheme was
331 related to the following binary outcomes: $D_{2cc, replanned}$ greater or less than the dose limit and model classification
332 accurate or inaccurate. All tests were insignificant, demonstrating that the variable prescriptions within our patient
333 cohort were not systematically impacting any of our key results.

334 One drawback of the proposed algorithm is that the necessity of needle supplementation is evaluated after
335 implantation of the applicator and imaging. This means that the model cannot guide needle decision-making for the
336 first fraction, and so the decision purely depends on the physician expertise, preference and fractionation schema.
337 However, predictions could still be used to inform needle implantation of subsequent fractions. In addition, our
338 current models do not provide guidance on the number and location of needles required to meet dose criteria.
339 Another limitation is that sensitivity and specificity of model classification was not tested on an independent dataset
340 and this may bias our results towards better model classification. Future work will examine the utility of this
341 decision-support tool and the standardized planning workflow in clinical practice.

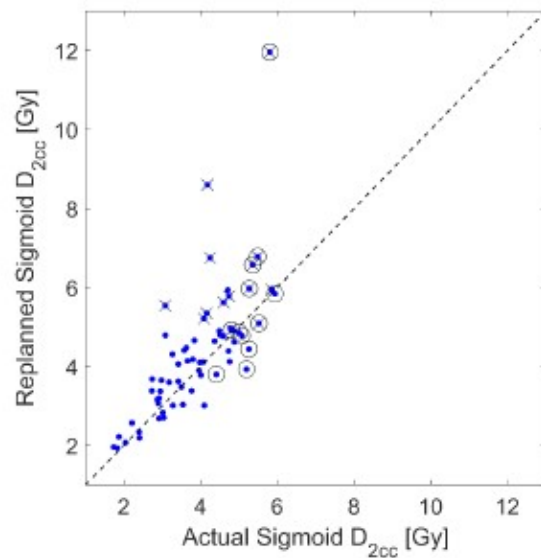
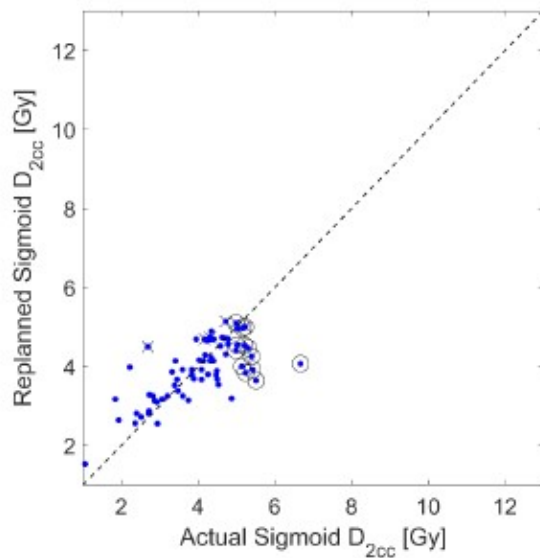
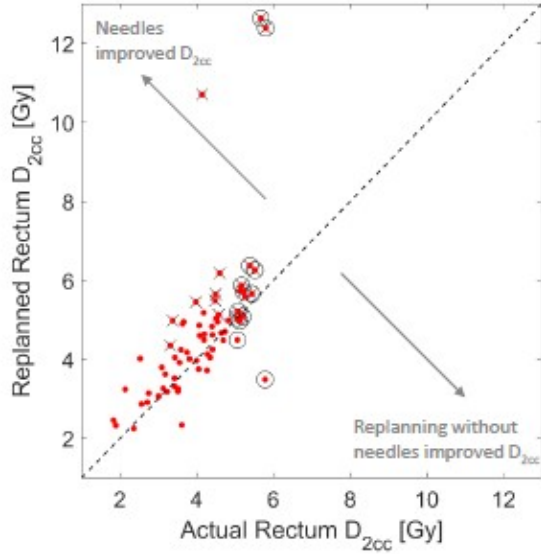
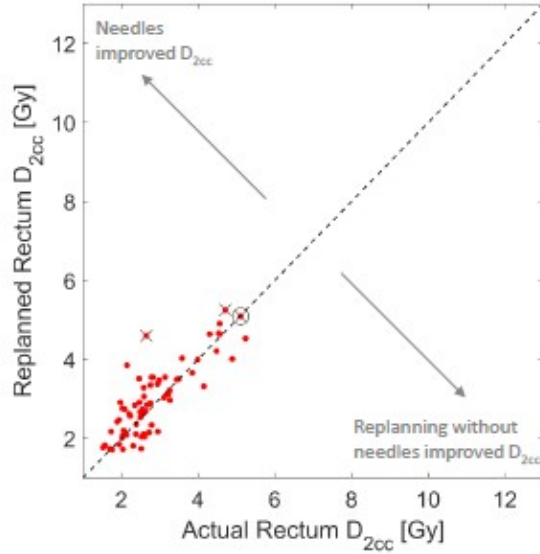
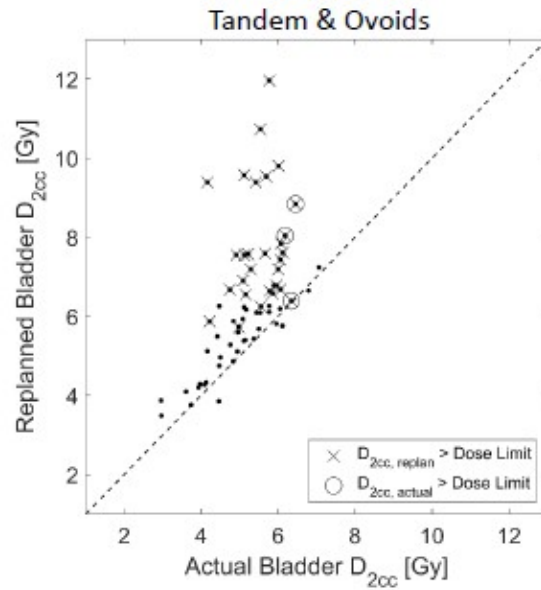
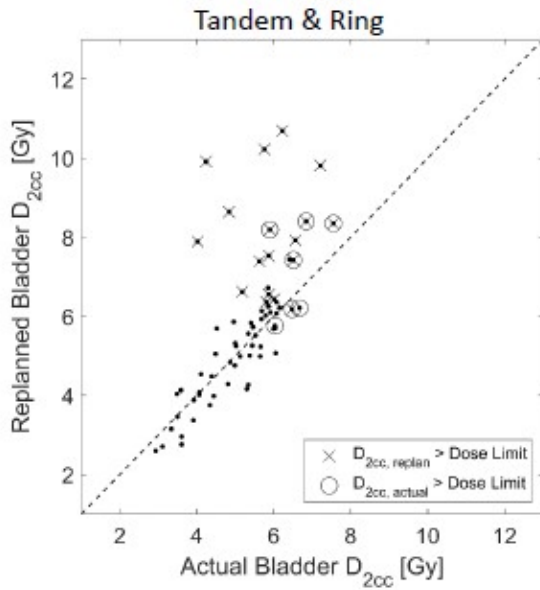
342 **Conclusion**

343 The benefit of knowledge-based intracavitary models to support the decision to use needles was
344 demonstrated. Standardized replanning of hybrid cases without needles confirmed model prediction accuracy. ROC
345 curves and AUCs demonstrated the discrimination accuracy of the tool, which featured much higher sensitivity and

346 specificity than our current clinical process for needle classification. The analysis showed that needles are sometimes
347 avoidable with little detriment to the patient, but could reduce organ dose, especially for the bladder. In summary,
348 standardized planning driven by knowledge-based dose predictions could influence needle usage, serve as guidelines
349 for physicians and decrease variations between treatment plans.

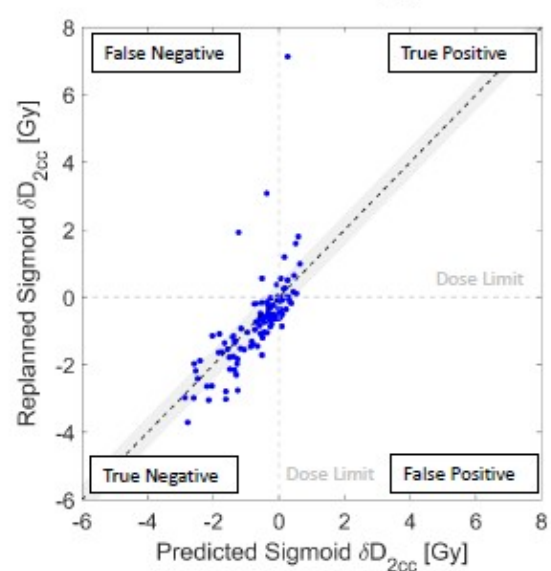
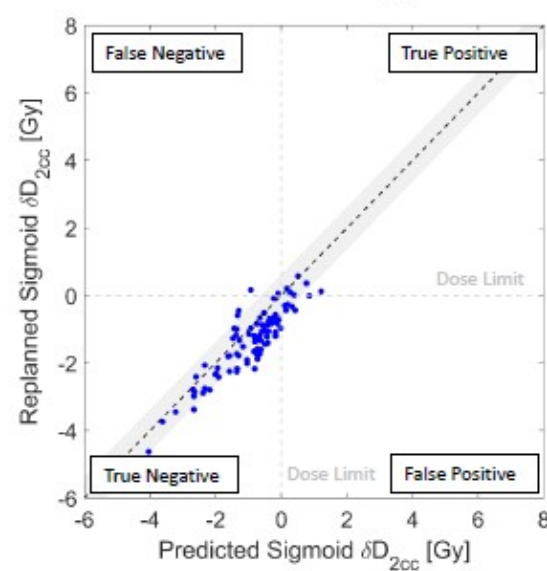
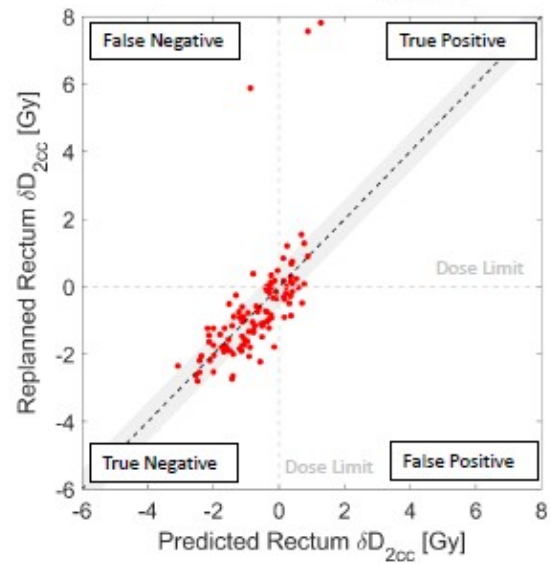
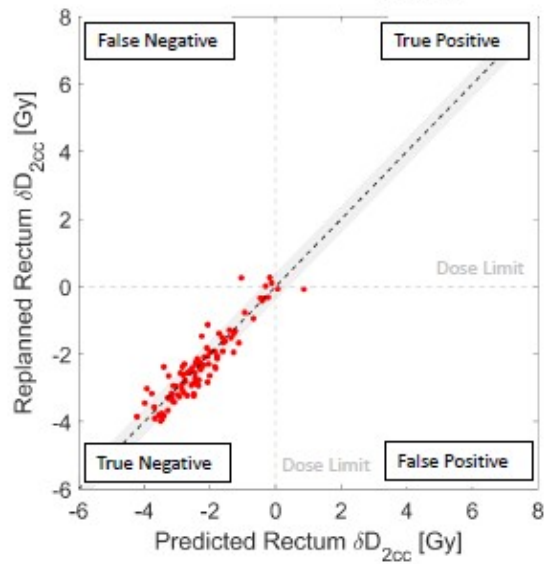
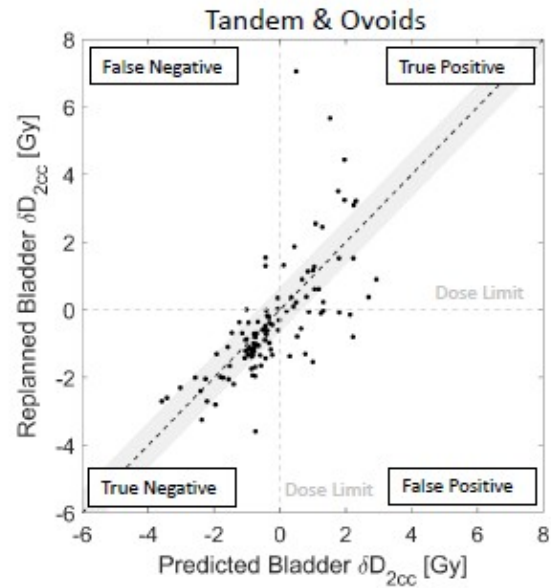
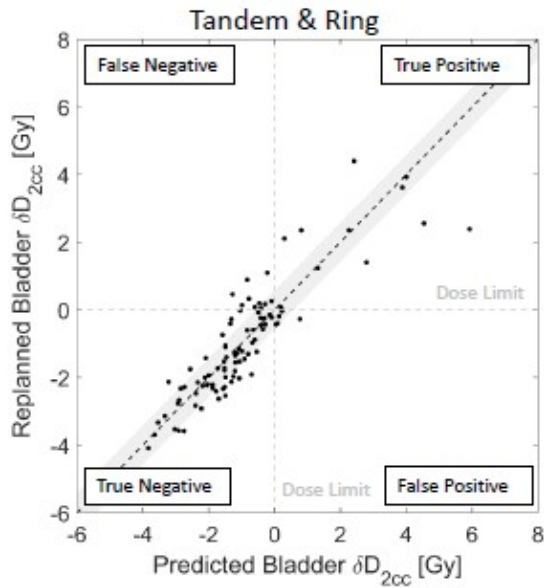
350 **Disclosure**

351 Dr. Meyers, Moore and Mayadev report grants from Padres Pedal the Cause during the conduct of the study. Dr.
352 Moore acknowledges funding support from AHRQ (R01 HS025440-01), has a patent Developing Predictive Dose-
353 Volume Relationships for a Radiotherapy Treatment licensed to Varian Medical Systems, and a patent for
354 knowledge-based prediction of three-dimensional dose distributions pending. Outside the submitted work, Dr.
355 Moore, Ray, Brown and Scanderbeg acknowledge research funding, travel support, and honoraria from Varian
356 Medical Systems; Dr. Scanderbeg reports personal research funding from Merit Medical; Dr. Mayadev reports
357 personal fees from AstraZeneca, grants from NRG Oncology and GOG Foundation, and personal fees from Varian
358 Medical Systems; Dr. Simon reports personal fees from Courage Health, Inc.



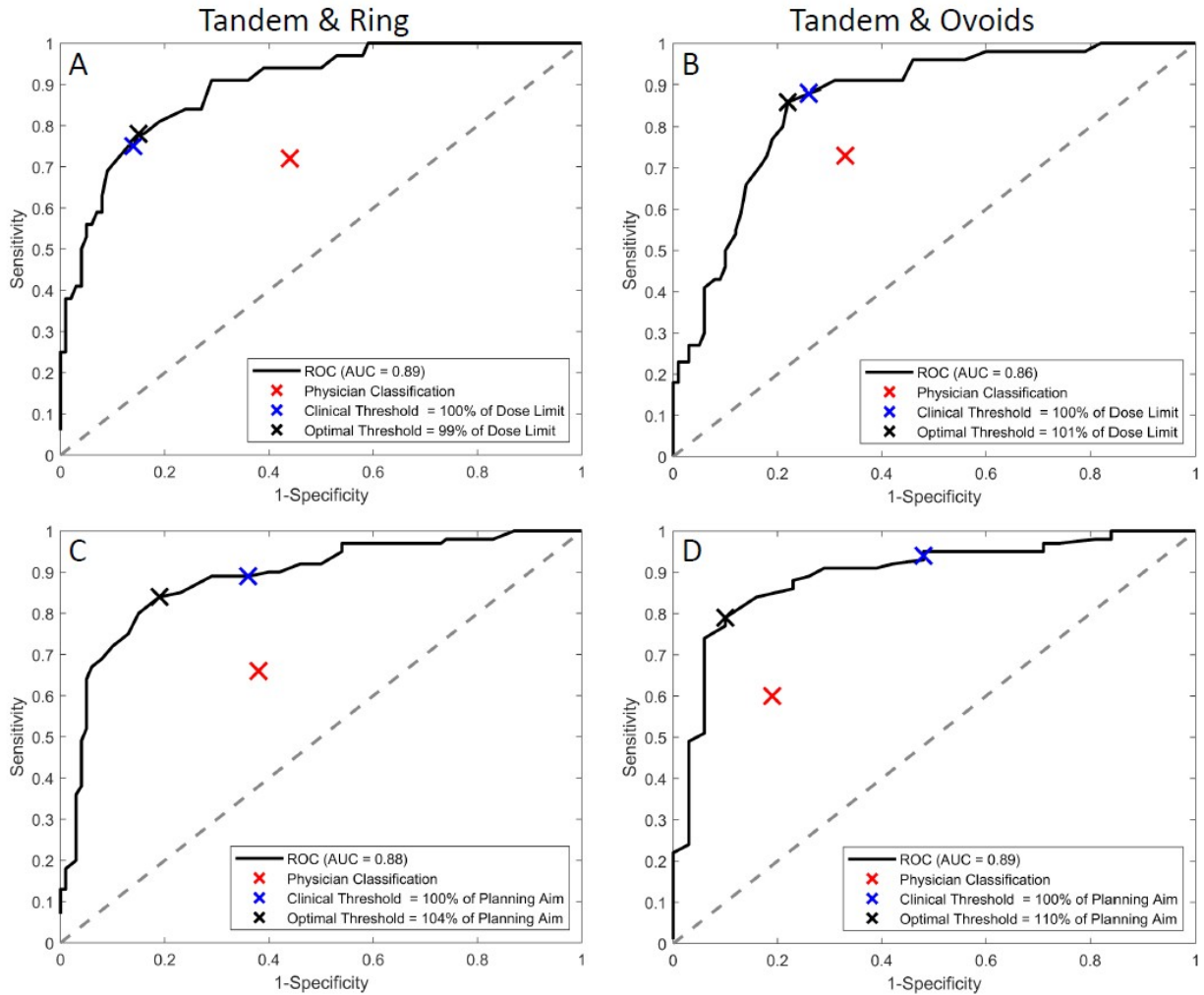
360 **Figure 1.** D_{2cc} values established by replanning the hybrid cases to purely intracavitary
361 treatment plans (i.e. organ dose without needles), compared to actual D_{2cc} values obtained
362 from clinical hybrid plans (i.e. organ dose with needles). Replans required HRCTV D90 to fall
363 between 85-90Gy EQD2. The dashed black line denotes the one-to-one line of replanned and
364 actual dose. Values below the dashed line show cases where replanning without needles
365 improved D_{2cc} values. Values above the dashed line show cases where the usage of needles
366 improved the D_{2cc} values; $x = D_{2cc, \text{replan}} > \text{dose limit}$; $o = D_{2cc, \text{actual}} > \text{dose limit}$.

367



369 **Figure 2.** Predicted deviations from dose limits compared to those of the replans, where
370 $\delta D_{2cc} = D_{2cc}$ -patient and organ-specific dose limit. Predictions were obtained from
371 knowledge-based intracavitary models, while replan values were established by re-planning
372 the hybrid cases without needles, requiring a HRCTV D90 of 85-90Gy EQD2. The dashed
373 black line denotes the one-to-one line of replanned and predicted doses, and the model
374 precision is displayed as the grey band. The four quadrants are labeled as if a threshold of
375 100% of the dose limits was used for model classification (although this threshold was
376 varied in the ROC analysis). Cases located in the bottom left and top right quadrant were
377 correctly predicted as not needing needles (true negatives) or needing needle
378 supplementation (true positives). Cases in the top left quadrant were falsely classified that
379 needles are unnecessary (false negatives), while cases in the bottom right quadrant were
380 predicted to require needles but could have been met the dose limits with the standard
381 applicator alone (false positives).

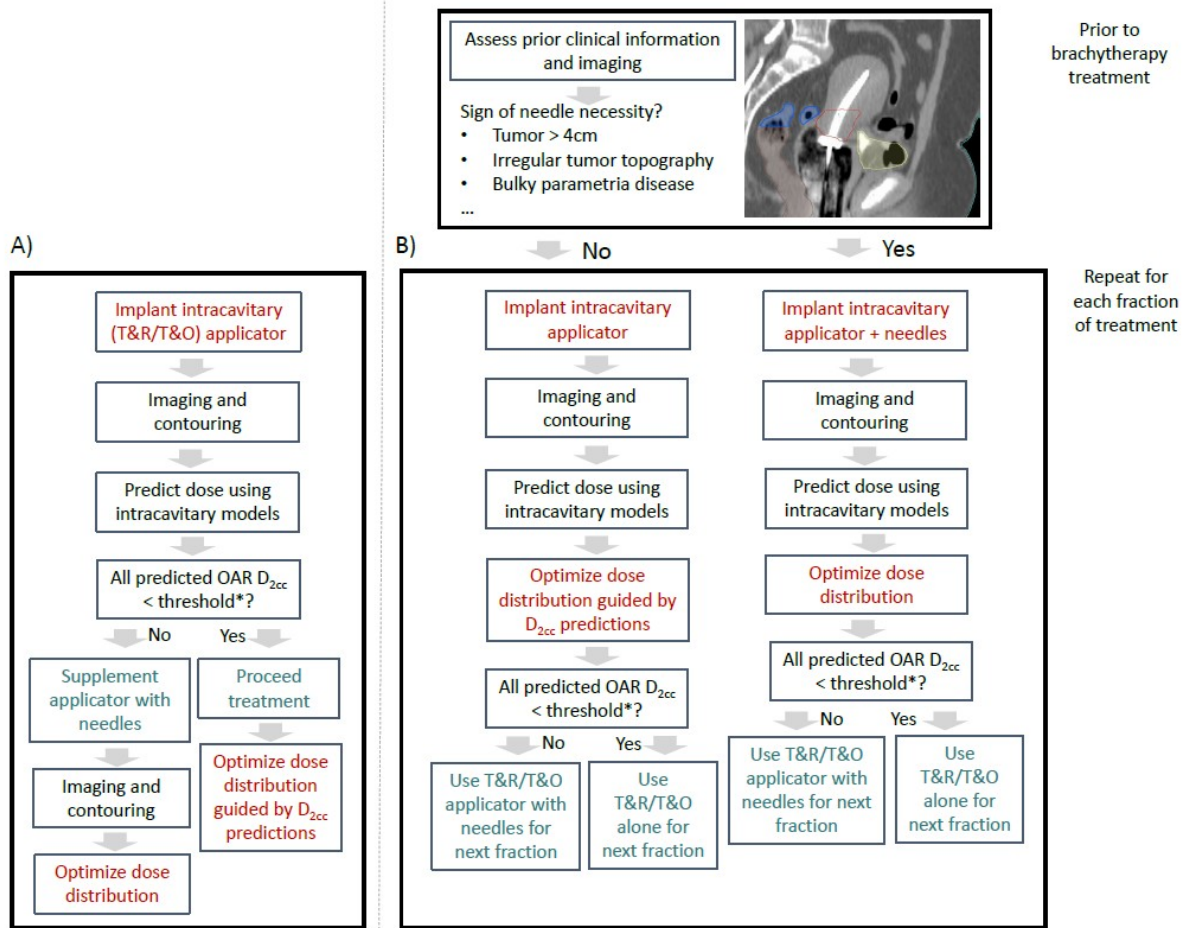
382



383

384 **Figure 3.** ROC curve and corresponding area-under-curve (AUC) for both applicators,
 385 considering all organ doses. Replanned D_{2cc} values are used as true condition (where any
 386 case with an organ dose exceeding the dose limit (A and B) or the planning aim (C and D) is
 387 considered to be a true needle case) and compared to the predicted D_{2cc} values for varying
 388 thresholds. The red cross indicates our current clinically achieved values of specificity and
 389 sensitivity classified by the physician; --- boundary marking the goodness of the prediction
 390 being better than randomness.

391



392

393 **Figure 4.** Suggested clinical workflows using the dose predictions to guide needle
 394 supplementation and treatment planning. (A) proposed workflow for centers where adding
 395 needles after imaging and subsequent re-imaging are feasible, such as those with image-
 396 guided brachytherapy suites. If this is not feasible, (B) shows an alternative workflow where
 397 predictions are used to guide decision-making for the subsequent fraction. *Optimal
 398 threshold defined by the maximal Youden index, i.e. 99% (101%) of dose limit or 104%
 399 (110%) of planning aim for T&R (T&O), where use of dose limits or planning aims is based on
 400 physician preference.

401

402 **Table 1.** Summary of patient specifications of evaluated hybrid cases and intracavitary model data; IC
 403 = intracavitary; IS = interstitial; T&R = tandem and ring applicator; T&O = tandem and ovoids
 404 applicator; T&RN = hybrid tandem and ring applicator with needles; T&ON = hybrid tandem and
 405 ovoids applicator with needles; HRCTV = high-risk clinical target volume.

Type	Parameter	Applicator	Specification	Value
Intracavitary	Number of included patients	T&R	Total	47
		T&O	Total	36
	Number of included fractions	T&R	Training cases	75
			Validation cases	38
		T&O	Training cases	80
			Validation cases	32
	HRCTV volume [cc]	T&R	Median (Range)	18.00 (4.9 – 40.2)
		T&O	Median (Range)	17.90.49 (7.7 - 65.7)
	Number of patients receiving each prescription (total number fractions x dose per fraction)	T&R	3 x 7.0Gy	2
			3 x 7.5Gy	2
3 x 8.0Gy			12	
3 x 8.5Gy			1	
4 x 5.5Gy			1	
4 x 7.0Gy			29	
4 x 7.5Gy			1	
5 x 6.0Gy			3	
T&O		3 x 8.0Gy	3	
		4 x 7.0Gy	15	
		4 x 6.0Gy	1	
		5 x 5.5Gy	4	
		5 x 6.0Gy	12	
		5 x 6.8Gy	1	
Number of patients with each tumor stage	T&R	I	22	
		II	19	
		III	5	
		IV	1	
	T&O	I	13	
		II	20	
		III	3	
		IV	0	
Hybrid (IC /IS)	Number of included patients	T&RN	Total	32
		T&ON	Total	28

Number of replanned fractions	T&RN	Total	70
	T&ON	Total	67
HRCTV volume [cc]	T&RN	Median (Range)	24.6 (9.7 – 73.4)
	T&ON	Median (Range)	30.60 (12.7 - 97.6)
Number of patients receiving each prescription (total number fractions x dose per fraction)	T&RN	3 x 7.0Gy	1
		3 x 8.0Gy	9
		4 x 7.0Gy	21
		5 x 6.0Gy	1
	T&ON	3 x 7.0Gy	1
		3 x 8.0Gy	4
		4 x 7.0Gy	19
		4 x 7.5Gy	2*
		5 x 5.0Gy	2
	Number of patients with each tumor stage	T&RN	I
II			15
III			7
IV			0
T&ON		I	9
		II	10
		III	7
		IV	0

406 *for one patient, one of the fractions was treated with 5Gy

407

408 **Table 2.** Results of the comparison of the predicted D_{2cc} values to the replanned D_{2cc} values, where the
409 replanned value represents the true result. Percentage of total amount of cases is shown where the
410 argument stated as parameter is met. OAR = organs-at-risk; T&R = tandem and ring applicator; T&O =
411 tandem and ovoids applicator; AUC = area under curve; σ_{Model} = model-prediction precision

Parameter	OAR	Result	
		T&R	T&O
Replanned $D_{2cc} < \text{OAR limit}$	Bladder	71%	55%
	Rectum	96%	70%
	Sigmoid	91%	78%
	All OARs	67%	39%
Replanned $D_{2cc} < \text{actual clinical } D_{2cc}$	Bladder	40%	6%
	Rectum	36%	28%
	Sigmoid	49%	34%
	All OARs	21%	1%
Replanned $D_{2cc} < \text{predicted } D_{2cc} + \sigma_{Model}$	Bladder	83%	79%
	Rectum	87%	90%
	Sigmoid	97%	82%
	All OARs	77%	68%
D_{2cc} prediction met OAR limits, but replanned $D_{2cc} > \text{OAR limit}$	Bladder	10 %	6%
	Rectum	4%	7%
	Sigmoid	3%	6%
	All OARs	9%	-
D_{2cc} prediction $> \text{OAR limits}$, but replanned $D_{2cc} < \text{OAR limit}$	Bladder	4%	18%
	Rectum	3%	13%
	Sigmoid	9%	10%
	All OARs	11%	19%
AUC of hybrid + intracavitary cases considering dose limit	Bladder	0.82	0.88
	Rectum	0.97	0.90
	Sigmoid	0.92	0.86
	All OARs	0.89	0.86
AUC of hybrid + intracavitary cases considering planning aims	Bladder	0.83	0.86
	Rectum	0.92	0.93
	Sigmoid	0.92	0.86
	All OARs	0.88	0.90

412

413

414 References

- 415 [1] Gill BS, Lin JF, Krivak TC, et al. National Cancer Data Base Analysis of Radiation Therapy Consolidation
416 Modality for Cervical Cancer: The Impact of New Technological Advancements. *Int J Radiat Oncol* 2014;
417 90: 1083–1090.
- 418 [2] Han K, Milosevic M, Fyles A, et al. Trends in the Utilization of Brachytherapy in Cervical Cancer in the
419 United States. *Int J Radiat Oncol Biol Phys* 2013; 87: 111–119.
- 420 [3] Viswanathan AN, Thomadsen B, American Brachytherapy Society Cervical Cancer Recommendations
421 Committee, et al. American Brachytherapy Society consensus guidelines for locally advanced carcinoma of
422 the cervix. Part I: general principles. *Brachytherapy* 2012; 11: 33–46.
- 423 [4] Lombe D, Crook J, Bachand F, et al. The addition of interstitial needles to intracavitary applicators in the
424 treatment of locally advanced cervical cancer: Why is this important and how to implement in low- and
425 middle-income countries? *Brachytherapy* 2020; 19: 316–322.
- 426 [5] Walter F, Maihöfer C, Schüttrumpf L, et al. Combined intracavitary and interstitial brachytherapy of cervical
427 cancer using the novel hybrid applicator Venezia: Clinical feasibility and initial results. *Brachytherapy* 2018;
428 17: 775–781.
- 429 [6] Nomden CN, de Leeuw AAC, Moerland MA, et al. Clinical Use of the Utrecht Applicator for Combined
430 Intracavitary/Interstitial Brachytherapy Treatment in Locally Advanced Cervical Cancer. *Int J Radiat Oncol*
431 2012; 82: 1424–1430.
- 432 [7] Dimopoulos JCA, Kirisits C, Petric P, et al. The Vienna applicator for combined intracavitary and interstitial
433 brachytherapy of cervical cancer: Clinical feasibility and preliminary results. *Int J Radiat Oncol* 2006; 66:
434 83–90.
- 435 [8] Appenzoller LM, Michalski JM, Thorstad WL, et al. Predicting dose-volume histograms for organs-at-risk
436 in IMRT planning. *Med Phys* 2012; 39: 7446–7461.
- 437 [9] Ge Y, Wu QJ. Knowledge-based planning for intensity-modulated radiation therapy: A review of data-
438 driven approaches. *Med Phys* 2019; 46: 2760–2775.
- 439 [10] Shiraishi S, Moore KL. Knowledge-based prediction of three-dimensional dose distributions for external
440 beam radiotherapy. *Med Phys* 2016; 43: 378–387.
- 441 [11] Nicolae A, Morton G, Chung H, et al. Evaluation of a Machine-Learning Algorithm for Treatment Planning
442 in Prostate Low-Dose-Rate Brachytherapy. *Int J Radiat Oncol* 2017; 97: 822–829.
- 443 [12] Shen C, Gonzalez Y, Klages P, et al. Intelligent inverse treatment planning via deep reinforcement learning,
444 a proof-of-principle study in high dose-rate brachytherapy for cervical cancer. *Phys Med Biol* 2019; 64:
445 115013.
- 446 [13] Yusufaly TI, Kallis K, Simon A, et al. A knowledge-based organ dose prediction tool for brachytherapy
447 treatment planning of patients with cervical cancer. *Brachytherapy* 2020; 19: 624–634.
- 448 [14] Cornell M, Kaderka R, Hild SJ, et al. Noninferiority Study of Automated Knowledge-Based Planning
449 Versus Human-Driven Optimization Across Multiple Disease Sites. *Int J Radiat Oncol Biol Phys* 2020; 106:
450 430–439.

- 451 [15] Tol JP, Dahele M, Gregoire V, et al. Analysis of EORTC-1219-DAHANCA-29 trial plans demonstrates the
452 potential of knowledge-based planning to provide patient-specific treatment plan quality assurance.
453 *Radiother Oncol* 2019; 130: 75–81.
- 454 [16] Younge KC, Marsh RB, Owen D, et al. Improving Quality and Consistency in NRG Oncology Radiation
455 Therapy Oncology Group 0631 for Spine Radiosurgery via Knowledge-Based Planning. *Int J Radiat Oncol*
456 *Biol Phys* 2018; 100: 1067–1074.
- 457 [17] Giaddui T, Geng H, Chen Q, et al. Offline Quality Assurance for Intensity Modulated Radiation Therapy
458 Treatment Plans for NRG-HN001 Head and Neck Clinical Trial Using Knowledge-Based Planning. *Adv*
459 *Radiat Oncol* 2020; 5: 1342–1349.
- 460 [18] Pötter R, Haie-Meder C, Limbergen EV, et al. Recommendations from gynaecological (GYN) GEC ESTRO
461 working group (II): Concepts and terms in 3D image-based treatment planning in cervix cancer
462 brachytherapy—3D dose volume parameters and aspects of 3D image-based anatomy, radiation physics,
463 radiobiology. *Radiother Oncol* 2006; 78: 67–77.
- 464 [19] Chino J, Annunziata CM, Beriwal S, et al. Radiation Therapy for Cervical Cancer: Executive Summary of
465 an ASTRO Clinical Practice Guideline. *Pract Radiat Oncol* 2020; 10: 220–234.
- 466 [20] Covele BM, Puri KS, Kallis K, et al. ORBIT-RT: A Real-Time, Open Platform for Knowledge-Based
467 Quality Control of Radiotherapy Treatment Planning. *JCO Clin Cancer Inform* 2021; 134–142.
- 468 [21] Biltekin F, Gultekin M, Yilmaz MT, et al. Dosimetric comparison of two different applicators and rectal
469 retraction methods used in inverse optimization-based intracavitary brachytherapy for cervical cancer. *J*
470 *Contemp Brachytherapy* 2020; 12: 35–43.
- 471 [22] Serban M, Kirisits C, de Leeuw A, et al. Ring Versus Ovoids and Intracavitary Versus Intracavitary-
472 Interstitial Applicators in Cervical Cancer Brachytherapy: Results From the EMBRACE I Study. *Int J*
473 *Radiat Oncol Biol Phys* 2020; 106: 1052–1062.
- 474 [23] Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. *Biom J Biom Z*
475 2005; 47: 458–472.
- 476 [24] Damato AL, Viswanathan AN, Don SM, et al. A system to use electromagnetic tracking for the quality
477 assurance of brachytherapy catheter digitization. *Med Phys* 2014; 41: 101702.
- 478 [25] Jung H, Shen C, Gonzalez Y, et al. Deep-learning assisted automatic digitization of interstitial needles in 3D
479 CT image based high dose-rate brachytherapy of gynecological cancer. *Phys Med Biol* 2019; 64: 215003.
- 480 [26] Zhou Y, Klages P, Tan J, et al. Automated high-dose rate brachytherapy treatment planning for a single-
481 channel vaginal cylinder applicator. *Phys Med Biol* 2017; 62: 4361–4374.
- 482 [27] Cv G, Kp A, R M, et al. Real-time inverse high-dose-rate brachytherapy planning with catheter optimization
483 by compressed sensing-inspired optimization strategies. *Phys Med Biol* 2016; 61: 5956–5972.
- 484 [28] Fokdal L, Sturdza A, Mazon R, et al. Image guided adaptive brachytherapy with combined intracavitary
485 and interstitial technique improves the therapeutic ratio in locally advanced cervical cancer: Analysis from
486 the retroEMBRACE study. *Radiother Oncol* 2016; 120: 434–440.
- 487 [29] Kaderka R, Mundt RC, Li N, et al. Automated Closed- and Open-Loop Validation of Knowledge-Based
488 Planning Routines Across Multiple Disease Sites. *Pract Radiat Oncol* 2019; 9: 257–265.

- 489 [30] Moore KL, Schmidt R, Moiseenko V, et al. Quantifying Unnecessary Normal Tissue Complication Risks
490 due to Suboptimal Planning: A Secondary Study of RTOG 0126. *Int J Radiat Oncol* 2015; 92: 228–235.
- 491 [31] Nelms BE, Chan MF, Jarry G, et al. Evaluating IMRT and VMAT dose accuracy: Practical examples of
492 failure to detect systematic errors when applying a commonly used metric and action levels. *Med Phys* 2013;
493 40: 111722.
- 494 [32] Tambas M, Tavli B, Bilici N, et al. Computed Tomography-Guided Optimization of Needle Insertion for
495 Combined Intracavitary/Interstitial Brachytherapy With Utrecht Applicator in Locally Advanced Cervical
496 Cancer. *Pract Radiat Oncol*. Epub ahead of print 5 February 2021. DOI: 10.1016/j.prro.2021.01.008.
- 497