A Differentiable Gaussian Mixture Model-based Scorecard for Cohort-Aware Radiotherapy Plan Evaluation and Optimization

Ho-Hsin Chang[1,2][0009−0005−2117−2148], Rex Cardan[1][0000−0001−9483−1398], Richard A. Popple[1][0000−0003−1252−738X], Dennis N. Stanley[1][0000−0003−4792−748X], John B. Fiveash[1][0000−0002−1129−2884], Sandeep Bodduluri[2][0000−0001−8201−5262], Joseph Harms[3][0000−0003−2441−9853], and Carlos E. Cardenas[1,2][0000−0003−1414−3849]

[1] [1]Department of Radiation Oncology, [2]Department of Biomedical Engineering, University of Alabama at Birmingham, Birmingham, AL, USA
[2] Department of Radiation Oncology, Washington University School of Medicine, St. Louis, MO, USA
{changh@uab.edu, cecardenas@uabmc.edu}

Abstract. Review of radiotherapy (RT) treatment plans often relies on subjective analysis by physicists and physicians. Qualitative step-wise metrics are non-differentiable as optimization objectives, posing hurdles to automated RT planning. While continuous plan quality scorecards are used for retrospective evaluation, their use as optimization objectives remains limited. This paper proposes a data-driven, differentiable scorecard based on a Gaussian Mixture Model cumulative distribution function (GMM-CDF) for cohort-aware radiotherapy plan evaluation.
Historical clinical prostate plans were categorized by prescription and target complexity, including 70 Gy hypofractionation and 40 Gy SBRT cohorts. A 3-component GMM mapped dosimetric parameters to continuous percentile-like scores. The naturally differentiable GMM-CDF yields probability density functions as exact derivatives, enabling fast gradient-based optimization. The GMM-CDF scorecard was evaluated via four experiments: an 80/20 train/test split, a target-complexity cohort assessment, a cross-prescription mismatch test, and scalar dose-scaling optimization using an L-BFGS-B solver.
Results demonstrated similar grade distributions between train and testing set. Cohort-specific scorecards produced distinct organ-at-risk (OAR) score distributions versus generalized scorecards, particularly across single-, dual-, and tri-target groups. In cross-prescription testing, OAR scores shifted substantially with mismatched prescription scorecards, whereas prescription-normalized target scores remained relatively stable. In optimization testing, average plan scores improved by 22.8 points.
The proposed GMM-CDF scorecard provides a continuous, differentiable approach of selected DVH-based plan quality metrics, supporting differentiable statistical scorecards as optimization-aware plan quality tools.

Keywords: Knowledge-based Planing · Treatment Plan Optimization · Machine Learning.

1 Introduction

Automated treatment planning relies heavily on inverse optimization algorithms guided by clinical dose-volume metrics [1, 7]. However, standard clinical goals are inherently step-wise and non-differentiable, relying mathematically on the Heaviside step function. Consequently, these metrics render the optimization landscape highly non-convex and disjointed [3]. Commercial optimization engines often rely on heuristic penalties or numerical approximations, which can complicate optimization and may contribute to convergence toward suboptimal solutions. Furthermore, rigid, static metrics fail to capture the continuous nature of plan quality and lack the context of an institution's historical performance.

To address the subjectivity and variability inherent in human planning, quantitative Plan Quality Metrics (PQMs) was introduced[2]. Foundational work demonstrated that even when provided with identical, explicit clinical objectives, plan quality varied wildly due to differences in planner skill [4]. Subjective preferences in trading off dose-volume constraints introduce biases, meaning some goals cannot be satisfied concurrently and making plans non-reproducible [5]. This variability is an ongoing, well-documented issue across an even broader population of modern planning systems. This realization catalyzed the development of continuous dosimetric scorecards to standardize evaluation. By fitting historical, population-level institutional data to probability distributions, modern scorecards map raw dose metrics to standardized percentile scores, allowing physicians to evaluate exactly how a plan ranks against historically accepted clinical practices [6].

Recently, these scorecard mechanisms have expanded beyond retrospective evaluation into the realm of automated planning [7]. Adjusted PQMs are increasingly utilized to curate and filter high-quality training libraries for Knowledge-Based Planning (KBP) models, ensuring that predictive algorithms learn only from optimal historical data by identifying and removing suboptimal outlier plans [8–10]. Additionally, quantitative quality scores have been deployed as reward functions in deep reinforcement learning (RL) and explainable AI models to guide artificial intelligence [11, 15]. Furthermore, highly generalizable deep learning frameworks, such as nnDoseNet, are increasingly being deployed to automate dose prediction across diverse clinical scenarios and modalities, reducing dose optimization time from weeks to minutes [12, 13]. Yet, despite this trajectory toward automation, these statistical metrics have remained largely disconnected from the core, deterministic gradient-based optimization process of commercial treatment planning systems.

In this paper, we propose a differentiable GMM-CDF scorecard for cohort-aware radiotherapy plan evaluation. Rather than generating treatment plans directly, this approach maps selected DVH-based plan quality metrics to continuous percentile-like scores and provides analytical gradients that may be evaluated in future optimization workflows. As a proof of concept, we apply the scorecard to retrospective clinical prostate treatment plans and assess its behavior across train/test validation, target-complexity stratification, cross-prescription scoring, and scalar dose-scaling optimization.

2 Materials and Methods

2.1 Dataset and Clinical Cohort Definitions

A retrospective database of clinical prostate cancer radiation therapy plans was used to construct and evaluate the scorecards . All plans included in this analysis were for intact prostate treatment; post-operative prostate cases were excluded. Plans were categorized by prescription dose and target complexity to account for differences in anatomical and dosimetric difficulty:

- **70Rx ALL (N=452):** All plans with a standard hypofractionated primary prescription of 70 Gy in 28 fractions. This cohort was further subdivided into three target-complexity groups:
  - *70RxSingle (N=251):* Simple plans treating a single target volume (prostate only).
  - *70RxDual (N=157):* Complex plans treating two distinct targets: a primary volume (70 Gy) and a secondary volume (typically the pelvic lymph nodes receiving 50.4 to 56 Gy).
  - *70RxTri (N=44):* Highly complex plans treating three distinct target structures (primarily the prostate, pelvic lymph nodes, and nodal lesions) alongside the primary 70 Gy prescription.
- **40Rx (N=281):** All plans utilize a 5-fraction prostate SBRT approach, delivering 36.25 Gy to the whole prostate with an integrated boost of 40 Gy to the primary target.

Plans were categorized as *Accepted* if they were clinically approved for treatment and as *Unaccepted* if they were generated during planning review but not selected for treatment delivery. Unaccepted plans in this dataset primarily represented borderline cases in which more stringent clinical judgment was applied, rather than catastrophic dosimetric failures.

2.2 DVH Metrics Selection

A standardized set of dose-volume histogram (DVH) metrics was used to construct the scorecards. Target metrics were selected to evaluate both target coverage and high-dose behavior, while OAR metrics were adopted from institutional plan-check criteria. To enable comparison across prescription groups, target metrics were evaluated as percentages of the prescription dose. All cohorts used the same target metrics: D90[%], D95%[%], D98[%], D0.01cc[%], and D0.03cc[%]. Target coverage metrics were treated as target-achievement metrics (larger values correspond to higher scores), and high-dose metrics were treated as hotspot-control metrics (lower valu es correspond to higher scores).

OAR metrics were stratified by prescription cohort to reflect dose-specific institutional criteria and were treated as sparing metrics (lower values correspond to higher scores):

- **70Rx Cohorts (Single, Dual, Tri):**

  - *Rectum:* Mean dose [Gy], V35Gy [%], V40Gy [%], V50Gy [%], V60Gy [cc], V70Gy [cc]
  - *Bladder:* V31Gy [%], V50Gy [%], V60Gy [%]
  - *Bowel:* Maximum dose [Gy], V54Gy [cc]
  - *Femoral Heads (L/R):* V50Gy [%]
- **40Rx Cohort:**
  - *Rectum:* V18.13Gy [%], V29Gy [%], V32.63Gy [%], V34.40Gy [cc], V38.06Gy [cc]
  - *Bladder:* V18.13Gy [%], V32.63Gy [%], V38.06Gy [cc]
  - *Urethra:* Maximum dose [Gy]
  - *Femoral Heads (L/R):* V20Gy [cc], Maximum dose [Gy]

Each metric was modeled independently. Missing metric values were handled by excluding the missing metric from the affected sub-score calculation while retaining the remaining available metrics for that plan.

2.3 Mathematical Formulation of the Objective Function

The proposed plan evaluation system utilizes a differentiable statistical scorecard based on institutional historical data. For each individual dose metric, x, a score $S(x)$ was defined using the cumulative distribution function (CDF) of a K-component Gaussian Mixture Model (GMM) [14]. In this implementation, a 3-component GMM was fit separately for each metric within each cohort using only the Accepted training population.

GMMs were fit using a custom Expectation-Maximization implementation with diagonal covariance, k-means++ initialization, 5 random initializations, a convergence tolerance of 1×10^{-6} on log-likelihood change, and a maximum of 200 iterations per initialization. A regularization term of 1×10^{-6} was added to each component variance to prevent numerical collapse.

For target structures, higher metric values were assigned higher scores:

$$S(x|\theta) = 100 \times \sum_{k=1}^{3} w_k \Phi\left(\frac{x - \mu_k}{\sigma_k}\right) \tag{1}$$

For OARs and hotspot-control metrics, lower metric values were assigned higher scores:

$$S(x|\theta) = 100 \times \left(1 - \sum_{k=1}^{3} w_k \Phi\left(\frac{x - \mu_k}{\sigma_k}\right)\right) \tag{2}$$

where $\Phi(\cdot)$ denotes the standard normal CDF, and $\theta = \{w_k, \mu_k, \sigma_k\}$ represents the weight, mean, and standard deviation of each component.

The aggregate plan score was defined as the equally weighted average of the mean target score and mean OAR score:

$$P_{Score} = \frac{1}{2}\left(\frac{1}{M_T}\sum \text{Target Scores} + \frac{1}{M_O}\sum \text{OAR Scores}\right) \tag{3}$$

where M_T and M_O represent the number of available target and OAR metrics. For optimization, the objective function was defined as the negative plan score ($J(\alpha) = -P_{Score}$).

2.4 Clinical Grading Thresholds

Based on the distribution of the Accepted training set, an objective 5-tier clinical grading system was established using the population mean (μ) and standard deviation (σ):

- **Great:** $P_{Score} > \mu + 2\sigma$
- **Good:** $P_{Score} \in [\mu + 2\sigma, \mu + 1\sigma]$
- **Acceptable:** $P_{Score} \in [\mu - 1\sigma, \mu + 1\sigma]$
- **Nearly Acceptable:** $P_{Score} \in [\mu - 2\sigma, \mu - 1\sigma]$
- **Unacceptable:** $P_{Score} < \mu - 2\sigma$

2.5 Analytical Gradients

The analytical derivative of each score with respect to the metric value x was calculated from the corresponding GMM probability density function (PDF):

$$\frac{\partial S}{\partial x} = \pm 100 \times \sum_{k=1}^{3} \frac{w_k}{\sigma_k} \phi\left(\frac{x - \mu_k}{\sigma_k}\right) \tag{4}$$

where $\phi(\cdot)$ denotes the standard normal PDF. The positive sign was used for target metrics, and the negative sign for OAR metrics.

3 Experiments and Results

3.1 Validation via 80/20 Split and Clinical Sensitivity

An 80/20 train/test split on Accepted plans confirmed consistent score distributions between training and independent testing cohorts. As detailed in Table 1, among historically accepted clinical plans, 63.0% to 75.0% of training and testing cases were assigned to the Acceptable category. Unaccepted plans showed variable grade distributions, indicating partial but incomplete separation from accepted plans.

Figure 1 displays selected scorecard distributions and cohort-specific grading thresholds. For the Bladder V31Gy [%] metric, the Acceptable-grade interval was 11.7 to 26.6 points for the 70RxSingle cohort, 45.1 to 63.9 points for the 70RxDual cohort, and 45.5 to 61.2 points for the 70RxTri cohort. The generalized 70Rx ALL scorecard produced an Accepted-grade interval of 15.6 to 54.3 points. For target metrics, the grading thresholds showed smaller differences across cohorts than those observed for selected OAR metrics.

3.2 Generalized vs. Complexity-Specific Cohort Shifts

Target-complexity experiments revealed that utilizing cohort-specific scorecards significantly shifted scoring evaluations (Table 2). For *70RxSingle* test set evaluated with a cohort-specific scorecard, the mean plan score decreased by 6.00 points, while *70RxDual* and *70RxTri* evaluations resulted in plan score increases of 6.92 and 7.12 points, respectively. These shifts indicate that generalized scorecards can inflate scores for simpler geometries and overly penalize complex cases.

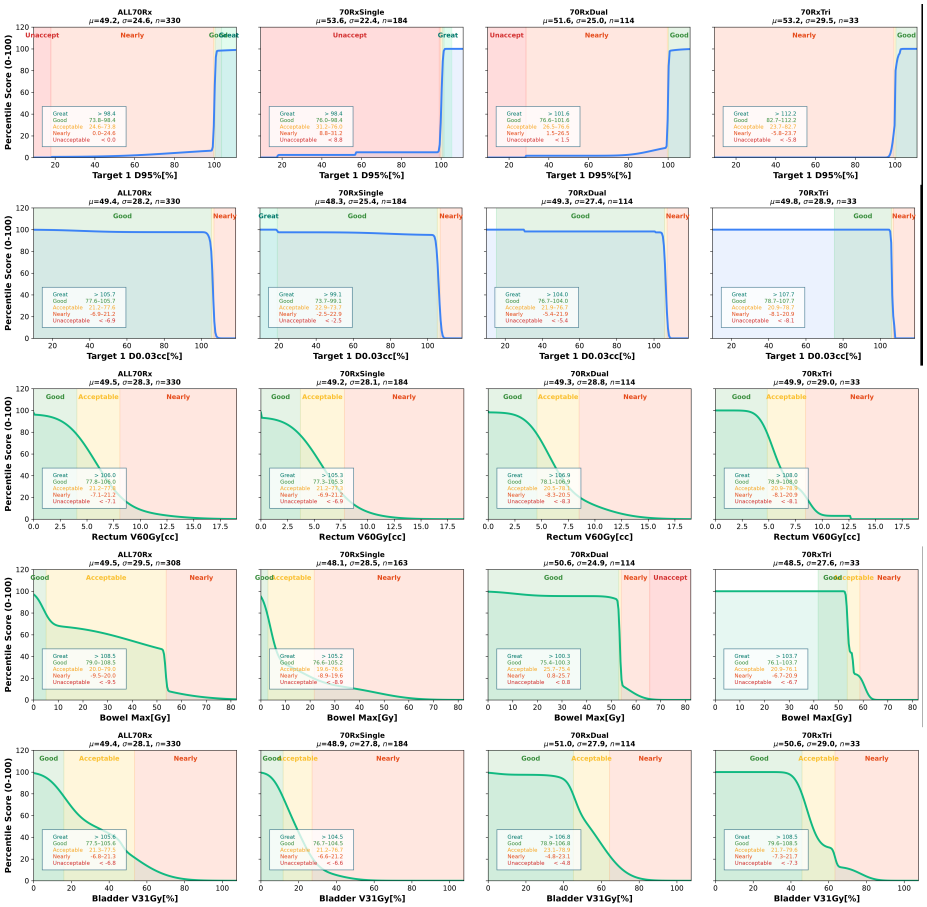

Fig. 1. Selected scorecard distribution. From left to right is 70Rx All, 70Rx Single, 70Rx Dual and 70Rx Tri, respectively. From top row to bottom represent different scorecard distribution of different metrics. D95%[%] of primary target (row 1) represent the target coverage, and D0.003cc[%] (row 2) represent hot spot control. The hot spot control metrics has scorecard direction like OARs. Row 3 to 5 represent different scorecard distribution of Rectum V60Gy[cc], Bowel Maximum [Gy] and Bladder V31Gy[%] respectively.

Table 1. Grade distributions for training, testing, and Unaccepted plans across scorecard cohorts. Training set and testing sets are clinically delivered plans, while unaccepted plans are rejected but considered nearly accepted cases.

Scorecard	Cohort Split	N	Great	Good	Acceptable	Nearly Acc.	Unacceptable
70Rx ALL	Train	330	11 (3%)	43 (13%)	220 (67%)	56 (17%)	0 (0%)
	Test	84	0 (0%)	12 (14%)	63 (75%)	8 (10%)	1 (1%)
	Unaccepted	38	8 (21%)	8 (21%)	16 (42%)	5 (13%)	1 (3%)
70Rx Single	Train	184	6 (3%)	25 (14%)	127 (69%)	22 (12%)	4 (2%)
	Test	46	3 (7%)	6 (13%)	30 (65%)	6 (13%)	1 (2%)
	Unaccepted	21	7 (33%)	6 (29%)	8 (38%)	0 (0%)	0 (0%)
70Rx Dual	Train	114	0 (0%)	16 (14%)	75 (66%)	20 (18%)	3 (3%)
	Test	29	0 (0%)	6 (21%)	20 (69%)	2 (7%)	1 (3%)
	Unaccepted	14	1 (7%)	0 (0%)	8 (57%)	4 (29%)	1 (7%)
70Rx Tri	Train	33	0 (0%)	9 (27%)	17 (52%)	7 (21%)	0 (0%)
	Test	8	0 (0%)	2 (25%)	6 (75%)	0 (0%)	0 (0%)
	Unaccepted	3	0 (0%)	1 (33%)	2 (67%)	0 (0%)	0 (0%)
40Rx	Train	223	3 (1%)	38 (17%)	147 (66%)	32 (14%)	3 (1%)
	Test	54	2 (4%)	9 (17%)	34 (63%)	8 (15%)	1 (2%)
	Unaccepted	4	0 (0%)	0 (0%)	2 (50%)	1 (25%)	1 (25%)

3.3 Cross-Prescription Scorecard Evaluation

Figure 2 summarizes the bidirectional cross-prescription evaluation. In cross-prescription evaluations, OAR scores shifted substantially when plans were evaluated using mismatched scorecards. When 70Rx plans were evaluated using the stricter 40Rx scorecard, the average plan score dropped sharply to 2.0. Conversely, evaluating 40Rx plans using the more relaxed 70Rx ALL OAR metrics yielded artificially inflated aggregate plan scores averaging 87.2 (corresponding to the Great range). Target scores, however, remained relatively stable across mismatched models due to their reliance on prescription-normalized metrics rather than absolute physical dose.

3.4 Scorecard-Guided Dose Optimization

A controlled scalar dose-scaling optimization was executed on 10 test cases using an L-BFGS-B optimizer (Table 3). The optimizer adjusted a single global parameter, α, initialized at α_0 (mean = 1.0167) to an optimized value α^* (mean = 1.0680). Optimizing the balanced plan score yielded a mean plan score increase of 22.8 points, driven by a mean target score increase of 47.0 points, alongside a modest mean OAR score decrease of 1.6 points. When the objective was restricted to target scores alone, 8 of 10 cases reached the upper scaling-factor boundary of $\alpha = 1.2$. When the objective was restricted to OAR scores alone, all 10 cases reached the lower scaling-factor boundary of $\alpha = 0.8$.

4 Discussion and Conclusion

This study demonstrates a cohort-aware, differentiable GMM-CDF scorecard that converts DVH metrics into continuous plan-quality scores. However, sev-

Table 2. Quantitative Complexity Shift Results on Testing Cases when switching from general scorecard (70Rx All) to specific scorecard (70Rx Single, Dual or Tri). Note: In the 70RxSingle testing set, 3 cases had incomplete target metric records, resulting in an active target scoring sample size of N=43. Plan and OAR metrics remained complete at N=46.

Cohort (Test)	Score	N	Score Shift	Better Grade	Worse Grade	Same Grade
70RxSingle	Plan	46	-6.00	0 (0.0%)	12 (26.1%)	34 (73.9%)
	Target	43	-3.27	0 (0.0%)	3 (6.7%)	42 (93.3%)
	OAR	46	-8.57	0 (0.0%)	20 (43.5%)	26 (56.5%)
70RxDual	Plan	29	+6.92	7 (24.1%)	0 (0.0%)	22 (75.9%)
	Target	29	+2.91	5 (17.2%)	0 (0.0%)	24 (82.8%)
	OAR	29	+10.94	13 (44.8%)	0 (0.0%)	16 (55.2%)
70RxTri	Plan	8	+7.12	3 (37.5%)	0 (0.0%)	5 (62.5%)
	Target	8	+1.35	3 (37.5%)	0 (0.0%)	5 (62.5%)
	OAR	8	+12.90	3 (37.5%)	0 (0.0%)	5 (62.5%)

Table 3. Scalar dose-scaling optimization results using the corresponding cohort-specific scorecard. Initialized at α_0 and optimized to α^*. Ten patient form different cohort are evaluated to test the ability of scorecard for optimization task .

Case	Cohort	Init α_0	Opt. α^*	Plan Score Shift	OAR Score Shift	Target Score Shift
001	40Rx	0.9914	1.0971	+9.8 (58.9 to 68.7)	+20.7 (54.3 to 75.0)	-1.2 (63.6 to 62.4)
002	40Rx	0.9978	1.1005	+24.1 (44.3 to 68.4)	+49.4 (25.7 to 75.1)	-1.1 (62.8 to 61.7)
003	40Rx	1.0322	1.1841	+16.7 (51.5 to 68.2)	+33.8 (41.0 to 74.8)	-0.5 (62.1 to 61.6)
004	70Rx Single	1.0128	1.0283	+25.2 (28.7 to 53.9)	+52.1 (22.1 to 74.2)	-1.5 (35.2 to 33.7)
005	70Rx Single	1.0046	1.0276	+20.5 (33.5 to 54.0)	+42.2 (32.6 to 74.8)	-1.2 (34.4 to 33.2)
006	70Rx Dual	1.0031	1.0396	+27.8 (28.0 to 55.8)	+56.8 (17.5 to 74.3)	-1.3 (38.5 to 37.2)
007	70Rx Dual	1.0079	1.0431	+27.4 (28.5 to 55.9)	+57.1 (16.6 to 73.7)	-2.2 (40.4 to 38.2)
008	70Rx Tri	1.0489	1.0952	+27.8 (23.5 to 51.3)	+58.9 (16.2 to 75.1)	-3.2 (30.7 to 27.5)
009	70Rx Tri	1.0077	1.0392	+25.0 (29.0 to 54.0)	+51.8 (22.4 to 74.2)	-1.8 (35.6 to 33.8)
010	70Rx Tri	1.0601	1.0257	+23.0 (31.6 to 54.6)	+47.3 (26.9 to 74.2)	-1.3 (36.3 to 35.0)
Avg.		**1.0167**	**1.0680**	**+22.8**	**+47.0**	**-1.6**

eral limitations remain.The gradient of a GMM-CDF score is proportional to its fitted probability density, the gradient can become weak when a severe metric violation lies in the far tail of the historical distribution.The Unaccepted plans used in this study do not exclusively represent poor DVH quality because plans may be rejected for non-dosimetric reasons; therefore, favorable scorecard grades among some Unaccepted plans reflect heterogeneous retrospective labels as well as imperfect score alignment. The scorecard's computational cost is highly efficient, scaling as $\mathcal{O}(M \times K)$ for M metrics and K mixture components, supporting its viability in complex optimization settings. However, the scalar-α experiment changes only global dose normalization and cannot redistribute dose or independently balance target and OAR objectives, so it demonstrates local score-function sensitivity rather than clinical inverse planning. Future work will relabel plans using a prospectively defined clinical cross-check checklist and clinician review to create a dosimetrically meaningful training reference. The scorecard will be combined with explicit clinical constraints and smooth differentiable penalty functions, such as softplus penalties, to maintain a stronger optimization signal for severe target and OAR violations. Finally, it will be

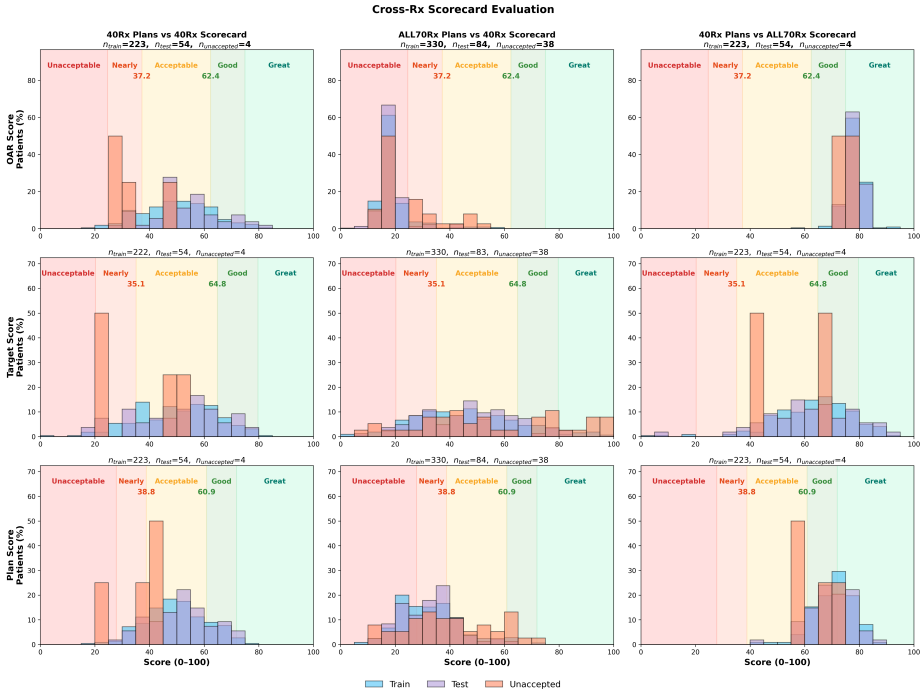

Fig. 2. Cross-prescription evaluation comparing target, OAR, and aggregate plan score distributions when plans were scored using matched and mismatched prescription-specific scorecards. From top to the bottom is different score of Plan, Target and OAR. From left to right is 40 Rx plans, evaluated by 40 Rx scorecard, 70 Rx plans evaluated on 40 Rx scorecard and 40 Rx plans evaluated on 70 Rx All scorecard.

evaluated within a multivariable, deliverable inverse-planning framework with a dose-influence model, spatial variables, and external validation.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

Ethical Statement IRB No.300012175 approved this retrospective study and waived informed consent for de-identified patient data.

References

1. Chui, C.S., Spirou, S.V.: Inverse planning algorithms for external beam radiation therapy. Med Dosim **26**(2), 189–197 (2001)
2. 1.McGarry, C. K. *et al.* The role of complexity metrics in a multi-institutional dosimetry audit of VMAT. *Br J Radiol* **89**, 20150445 (2016).
3. Maass, K., Kim, M., Aravkin, A.: A Nonconvex Optimization Approach to IMRT Planning with Dose-Volume Constraints. INFORMS Journal on Computing **34**(3), 1366–1386 (2022)

4. Nelms, B.E., Robinson, G., Markham, J., et al.: Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems. Pract Radiat Oncol **2**(4), 4 (2012)
5. Roy, A., Das, I.J., Nohadani, O.: On correlations in IMRT planning aims. J Appl Clin Med Phys **17**(6), 44–59 (2016)
6. Leone, A.O., Mohamed, A.S.R., Fuller, C.D., et al.: A Visualization and Radiation Treatment Plan Quality Scoring Method for Triage in a Population-Based Context. Adv Radiat Oncol **9**(8), 101533 (2024)
7. Nguyen, D., Lin, M.H., Sher, D., Lu, W., Jia, X., Jiang, S.: Advances in Automated Treatment Planning. Semin Radiat Oncol **32**(4), 343–350 (2022)
8. Tol, J.P., Delaney, A.R., Dahele, M., Slotman, B.J., Verbakel, W.F.A.R.: Evaluation of a knowledge-based planning solution for head and neck cancer. Int J Radiat Oncol Biol Phys **91**(3), 612–620 (2015)
9. Cardenas, C.E., Cardan, R.A., Harms, J., Simiele, E., Popple, R.A.: Knowledge-based planning, multicriteria optimization, and plan scorecards: A winning combination. Radiotherapy and Oncology **202**, 110598 (2025)
10. Delaney, A.R., Tol, J.P., Dahele, M., Cuijpers, J., Slotman, B.J., Verbakel, W.F.A.R.: Effect of Dosimetric Outliers on the Performance of a Commercial Knowledge-Based Planning Solution. Int J Radiat Oncol Biol Phys **94**(3), 469–477 (2016)
11. Shen, C., Gonzalez, Y., Klages, P., et al.: Intelligent inverse treatment planning via deep reinforcement learning, a proof-of-principle study in high dose-rate brachytherapy for cervical cancer. Phys Med Biol **64**(11), 115013 (2019)
12. Chang, H., Harms, J., Cardan, R.A., Fiveash, J.B., Popple, R.A., Cardenas, C.E.: nnDoseNet: Intuitive and flexible deep learning framework to train and evaluate radiotherapy dose prediction models. Computers in Biology and Medicine **198**, 111237 (2025)
13. Chang, H., Cardan, R.A., Nedunoori, R., et al.: Generalizable Deep Learning Framework for Radiotherapy Dose Prediction Across Cancer Sites, Prescriptions and Treatment Modalities. medRxiv (2026)
14. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum Likelihood from Incomplete Data Via the EM Algorithm. Royal Statistical Society Journal Series B: Methodological **39**(1), 1–22 (1977)
15. Ebrahimi, S., Lim, G.J.: A reinforcement learning approach for finding optimal policy of adaptive radiation therapy considering uncertain tumor biological response. Artif Intell Med **121**, 102193 (2021)