# Iterative Refinement of Radiation Therapy Dose Distribution Prediction for Accelerated Partial Breast Irradiation via Plan Scoring

**Ledi Wang, Rafe McBeth**

University of Pennsylvania

ledi.wang@pennmedicine.upenn.edu, rafe.mcbeth@pennmedicine.upenn.edu

## Abstract

The integration of artificial intelligence into clinical workflows offers substantial opportunities to enhance treatment quality in radiation oncology but also require methods that address the complexity of clinical decision making. Recent advances in dose prediction demonstrate the value of conditioning on patient specific anatomy. However, the tendency of deep learning models to regress toward cohort averages motivates strategies that preferentially learn from higher quality exemplars. Building on these insights, we present a framework that uses clinically relevant metrics to refine prediction quality for accelerated partial breast irradiation. A linear piecewise scoring system assigns normalized scores to each plan across dose volume histogram metrics covering target coverage and sparing of organs at risk. We applied this framework to a retrospective cohort of 550 patients treated at our institution, and we trained a three dimensional dose prediction model based on a hierarchically dense U Net. Preliminary results reveal wide variation in score distributions across the cohort, reflecting both patient specific complexity and variation in planning quality. We trained a baseline model and then applied our iterative refinement framework, which ranks cases by the composite quality score and, at each round, retains the highest scoring half for retraining. Models refined in this manner demonstrated promising improvements in predicted plan quality, supporting the effectiveness of this approach. These findings illustrate how scoring guided curation can align model behavior with clinical priorities and provide a path toward adaptive, high precision treatment planning tools.

## Introduction

Radiation oncology is well positioned to benefit from advances in artificial intelligence. Its end to end digital workflow, spanning imaging, contouring, inverse planning, and quality assurance, culminates in a computable treatment plan that directly parameterizes delivery on computer controlled linear accelerators and other image guided platforms.

This mature and data rich infrastructure, together with established standards for data exchange, verification, and audit, provides a strong basis for integrating artificial intelligence into treatment planning to improve efficiency, consistency, and overall plan quality.

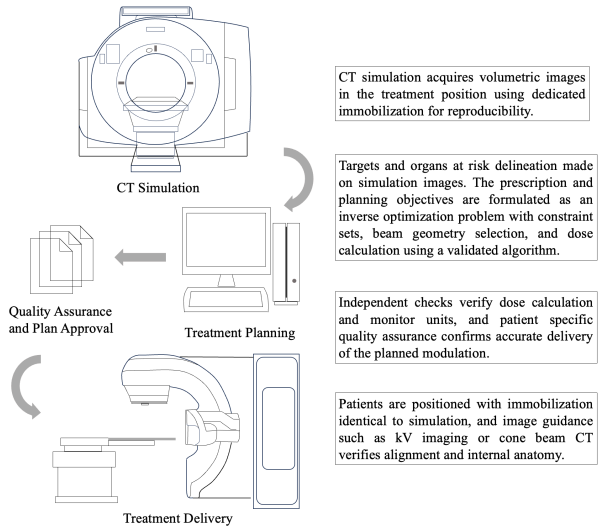


Figure 1: **Four Stage Radiation Oncology Workflow.** CT simulation acquires the reference anatomy, treatment planning converts clinical intent into a deliverable plan, quality assurance and plan approval confirm safety and accuracy, and treatment delivery executes the plan with image guidance and motion management.

At the same time, radiation oncology faces substantial workforce constraints, including a widening shortage of qualified specialists. These pressures are increasing demand for technologies that can reduce the workload on clinical teams while sustaining, and ideally improving, the quality of care.

Treatment planning in radiation oncology exhibits substantial variability, both between observers and within the same observer over time. Patient anatomy also varies widely, which complicates the evaluation of plan quality

across cases. Clinical scorecards are widely adopted to enforce adherence to target volume coverage and organs at risk (OAR) constraints, thereby establishing minimum standards for quality and safety. However, these tools provide limited granularity for robust comparative assessment across patients and are not designed to adapt or learn.

Meaningful variation in plan quality persists even when all plans satisfy clinical acceptability criteria. Because data driven dose prediction models tend to approximate the average quality of their training cohorts, we posit that curating the dataset toward higher quality exemplars will steer predictions toward superior plans with minimal loss of generalizability. To evaluate this, we developed an iterative framework guided by dynamic scoring functions that quantifies plan quality.

An important consideration is that some patients may achieve lower scores because of intrinsic anatomical complexity. Addressing this case mix variability is a priority in our ongoing work. We aim to develop methods that explicitly model patient specific anatomical complexity and adjust quality assessment and model refinement accordingly.

# Background

## Accelerated Partial Breast Irradiation

Breast-conserving therapy (BCT) traditionally couples lumpectomy with whole-breast irradiation (WBI), typically ~50 Gy in 25 fractions over 5–6 weeks, which lowers local failure and improves disease-specific outcomes with acceptable toxicity and cosmesis. Yet the length and logistics of WBI can deter some patients from pursuing BCT. APBI was developed to shorten adjuvant radiation by limiting treatment to the tumor bed with margin, thereby reducing treatment burden while attempting to preserve tumor control and minimize normal-tissue effects (Wang and Vicini 2005).

Several delivery strategies are in routine use. Multi-catheter interstitial brachytherapy has the longest follow up and enables conformal coverage through a catheter array. Balloon based brachytherapy and hybrid applicators broaden applicator based options. External beam approaches such as three dimensional conformal radiotherapy and intensity modulated radiotherapy avoid re-entry and can improve dose uniformity, while intraoperative radiotherapy offers single session delivery at the time of surgery but concentrates all target definition and quality assurance into one encounter (Njeh et. al. 2010).

Planning and delivery for external beam radiation therapy (EBRT) require attention to motion, setup reproducibility, and cavity delineation. Respiration and daily variation can motivate larger margins, with the tradeoff of more normal breast at low and mid dose; prone positioning can reduce motion and improve heart and lung sparing, though it requires dedicated immobilization and is not universally adopted. Delineation of the cavity varies among observers; surgical clips help reduce this variability. External beam tends to produce better dose homogeneity, which may support favorable cosmesis, but multiple field arrangements can increase integral dose to surrounding tissue (Offersen 2009).

## Radiation Therapy Dose Distribution Prediction

Radiation therapy dose distribution prediction aims to estimate a clinically achievable three-dimensional dose before full optimization, so that physician directives and plan generation can be guided by realistic targets. Early knowledge-based planning approaches learned dose–volume relationships from historical plans but were limited to one- or two-dimensional outputs such as dose volume histograms (DVHs) and depended on hand-crafted features, leaving spatial dose patterns to the planner's judgment. Deep learning enabled volumetric dose prediction directly from anatomy, and learns voxel-to-voxel mappings and support truly three-dimensional predictions within the planning workflow. These models can inform physician directive generation, plan generation, and quality assurance (Nguyen et. al. 2022).

Across the recent literature, convolutional networks, especially U-Net variants, dominate, with inputs typically comprising CT plus target and OAR contours, and outputs ranging from full 3D dose to DVHs. A comprehensive review identified 89 studies between 2018 and early 2024, with U-Net the most widely used architecture and ResNet the next most common, reflecting their strength in capturing local and global spatial hierarchies important for dose prediction (Kazemzadeh et. al. 2025).

Methodological breadth for dose prediction is growing. Multi-task, multimodal frameworks that learn across delivery platforms can preserve accuracy across systems and workflows. In one study spanning CyberKnife, MR-Linac (Unity), Ethos, TrueBeam, and GammaPod, multi-task learning reduced mean absolute percent error compared with single-task models and maintained competitive gamma pass rates, suggesting improved robustness without sacrificing dosimetric fidelity (Maniscalco et. al. 2024).

Translating prediction into clinical value requires staged integration, interpretability, and uncertainty quantification. Recommended steps include tool readiness checks, workflow and UI design, initial implementation with debugging, departmental approval, clinical deployment, and ongoing QA; uncertainty estimates and attention maps can help communicate model confidence and focus, though interpretation remains clinician-led. Within the broader trial and practice landscape, AI now predicts doses and fluences, supports plan validation and QA, and is being incorporated into adaptive and clinical-trial workflows, with attention to data curation, robustness, and bias mitigation (Xiao et. al. 2025).

### Knowledge Based Planning and Deep Learning Based Planning

Knowledge-based planning (KBP) arose as a data-driven strategy to improve the quality and efficiency of intensity modulated radiation therapy and volumetric modulated arc therapy (VMAT) by learning from prior clinical plans. In KBP, models exploit relationships between patient geometry, historical dosimetry, and planning parameters to predict achievable dose–volume goals or even voxel-level dose, thereby guiding or partially automating optimization. Methodologically, KBP is commonly organized into two families. Case and atlas-based methods retrieve anatomically similar exemplars using direct or indirect similarity measures and transfer "knowledge" ranging from beam geometry and objective weights to DVH constraints or voxel dose. By contrast, statistical and machine-learning models fit predictive functions—often multivariate regressions or ensemble learners—using handcrafted geometric and dosimetric features (Ge and Wu 2019).

Deep-learning-based planning extends this paradigm by learning direct mappings from patient anatomy to three-dimensional dose distributions, most often with encoder–decoder convolutional networks such as U-Net variants. These models have demonstrated accurate voxel-level dose prediction and can accommodate heterogeneous beam configurations, supporting broader clinical applicability.[9] A complementary line of work predicts dose distributions by conditioning on physician-specified DVH targets, which enables steering toward patient-specific priorities. Together they provide dose priors that can inform physician directives and accelerate plan generation (Ma et. al. 2021).

### Quality Measurement of Radiation Therapy Treatment Plan

Plan quality is best understood as the clinical suitability of the dose that is realistically deliverable, not simply the appearance of the nominal plan on the screen. It follows that quality reflects three intertwined elements: the dosimetric profile, the plan's robustness against uncertainties, and the complexity of delivery. The literature highlights that there is not yet a single agreed metric, and calls for tools that standardize evaluation, enable robustness analysis, and report complexity in routine planning (Hernandez et. al. 2020).

Quality measurement is confounded by variation in prescription and calculation conventions. Choice of algorithm and calculation settings also alters displayed isodoses and dose volume metrics (Hansen et. al. 2020).

Robustness measurement asks how the plan behaves across plausible setup and anatomy scenarios, either in worst case or probabilistic formulations. Evidence shows that scenario doses can diverge meaningfully from the approved nominal dose, arguing that robustness analysis should complement or replace margin only evaluation. What level of variation in target coverage and organ at risk dose should be deemed acceptable in routine reporting.

Complexity measurement captures how modulation and machine parameters influence deliverability and uncertainty. Greater complexity can degrade calculation accuracy and delivery stability, and is often uncorrelated with dosimetric quality, motivating explicit reporting and, when possible, regularization during optimization.

## Methods

### Data & Pre-Processing

The study cohort comprised 550 consecutive APBI VMAT cases treated at our institution. The dataset captures broad anatomic variability in breast size, laterality, tumor bed location, and spatial relationships between targets and OARs. All plans were evaluated using an adapted clinical scoring framework that translates institutional objectives into DVH metrics organized into priority tiers. For each case, a composite plan quality score was computed to summarize adherence to mandatory constraints and performance on lower priority objectives, enabling consistent comparison across patients.

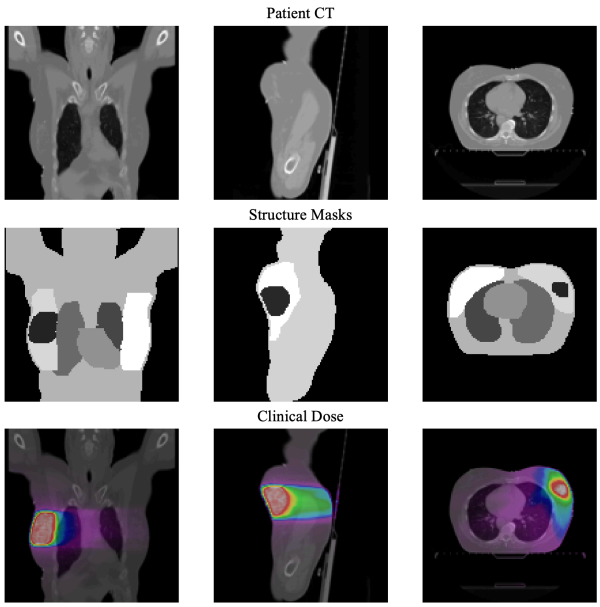


Figure 2: **Clinical data representation for APBI modeling: CT, structures, and dose.** Example case from the study cohort. Top panel shows the planning CT in treatment position. Middle panel overlays structure masks used as model inputs. Bottom panel displays the clinical plan dose as a color wash normalized to prescription with isodose contours. Images are axial. All masks and dose are resampled to the planning grid, and the set of CT plus masks forms the input channels, with dose serving as the prediction target.

## Translating Clinical Scorecards

Our EBRT APBI clinical scorecard operationalizes plan evaluation through DVH metrics organized by priority level. The primary requirement is target coverage, specified as 95% dose to planning target volume (PTV) achieve at least 28.5 Gy. Secondary priorities constrain focal hot spots in the external patient with dose to 0.03 cc at most 31.5 Gy and limit cardiac exposure with percent volume of heart receiving 3 Gy at most 7% and mean dose at most 1 Gy. They also restrict low and mid dose spillage to the ipsilateral breast with volume receiving 15 Gy at most 50% and 28.5 Gy at most 25%, and limit lung irradiation with ipsilateral lung volume receiving 10 Gy at most 15% and 5 Gy at most 50%. Tertiary items protect contralateral structures with contralateral breast dose to 0.03 cc at most 1 Gy and contralateral lung volume receiving 5 Gy at most 10%. Each objective is paired with an associated variation value to support consistent adjudication and composite scoring across patients.

| Structure Name | DVH Objective | Evaluator | Variation | Priority |
|---|---|---|---|---|
| PTVp_300 | D95%[Gy] | >=28.5 | 27 | 1 |
| External_Pat | D0.03cc[Gy] | <=31.5 | 33 | 2 |
| Heart | V3Gy[%] | <=7 | 10 | 2 |
| Heart | Mean[Gy] | <=1 | 2 | 2 |
| Breast_R | V15Gy[%] | <=50 | 55 | 2 |
| Breast_R | V28.5Gy[%] | <=25 | 30 | 2 |
| Breast_L | D0.03cc[Gy] | <=1 | 3 | 3 |
| Lung_R | V10Gy[%] | <=15 | 20 | 2 |
| Lung_R | V5Gy[%] | <=50 | 55 | 2 |
| Lung_L | V5Gy[%] | <=10 | 15 | 3 |

**Table 1: DVH-based Clinical Scorecard.** The scorecard is laterality aware. For right sided disease, ipsilateral structures map to Breast_R and Lung_R and contralateral structures map to Breast_L and Lung_L. For left sided disease, these assignments are reversed. Heart and External_Pat objectives apply in all cases, and target objectives for PTVp_300 are unchanged.

Beyond threshold checks, the composite score includes a penalty when the maximum dose exceeds 105% of the prescription. This discourages reliance on focal hot spots while preserving flexibility for anatomically challenging cases.

For each OAR and TV with their corresponding DVH metric, we define a piece wise linear scoring function $S(x)$ composed of $n$ linear segments:

$$S(x) = \begin{cases} f_1(x) = m_1 x + b, & x_{min,1} \le x \le x_{max,1} \\ f_2(x) = m_2 x + b, & x_{min,2} \le x \le x_{max,2} \\ \dots & \\ f_n(x) = m_n x + b, & x_{min,n} \le x \le x_{max,n} \end{cases}$$

Where $x$ represents the DVH metric value (dose or volume), $m_i$ and $b_i$ are the slope and $y$-intercept of the $i$-th linear segment, and $[x_{min,i}, x_{max,i}]$ defines the domain of the $i$-th segment.

The piecewise linear function is defined by a set of control points:

$$P = (x_1, y_1), (x_2, y_2), \dots, (x_{n+1}, y_{n+1})$$

For each linear segments.

$$m_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}, \qquad b_i = y_i - m_i \cdot x_i$$

To handle values outside the defined range:

$$S(x) = \begin{cases} S(x_{min}) = f_1(x_{min}), & x < x_{min} \\ f_i(x), & x_{min,i} \le x \le x_{max,i} \\ S(x_{max}) = f_n(x_{max}), & x > x_{max} \end{cases}$$

This ensures score clamping at boundary values, preventing unrealistic scores for extreme dosimetric values.

The volume receiving at least dose $x$ is calculated as:

$$V_x = \frac{|x \in \Omega_s : D(v) \ge x|}{|\Omega_s|} \times 100$$

Where $\Omega_s$ represents the structure volume, $D(v)$ is the dose at voxel $v$, and $|\cdot|$, denotes cardinality (voxel count). The dose delivered to $y$ of the structure volume:

$$D_y = \inf d : V_d \ge y$$

The average dose across the structure:

$$\overline{D} = \frac{1}{|\Omega_s|} \sum_{v \in \Omega_s} D(v)$$

Using the linear piecewise scoring function, we assigned normalized scores to each treatment plan across a variety of DVH metrics, including target coverage and OAR sparing criteria. For APBI patients, metrics were evaluated separately for ipsilateral and contralateral structures depending on the location of the seroma. The scoring framework incorporated bother evaluation goals and acceptable variation thresholds, with priority levels explicitly tied to clinical importance.

Priority 1 metrics, such as target coverage (e.g., D95% to the PTV), were weighted heavily to ensure no compromise in clinical safety or efficacy. Metrics with lower priority, such as dose to non-critical structures, were assigned milder

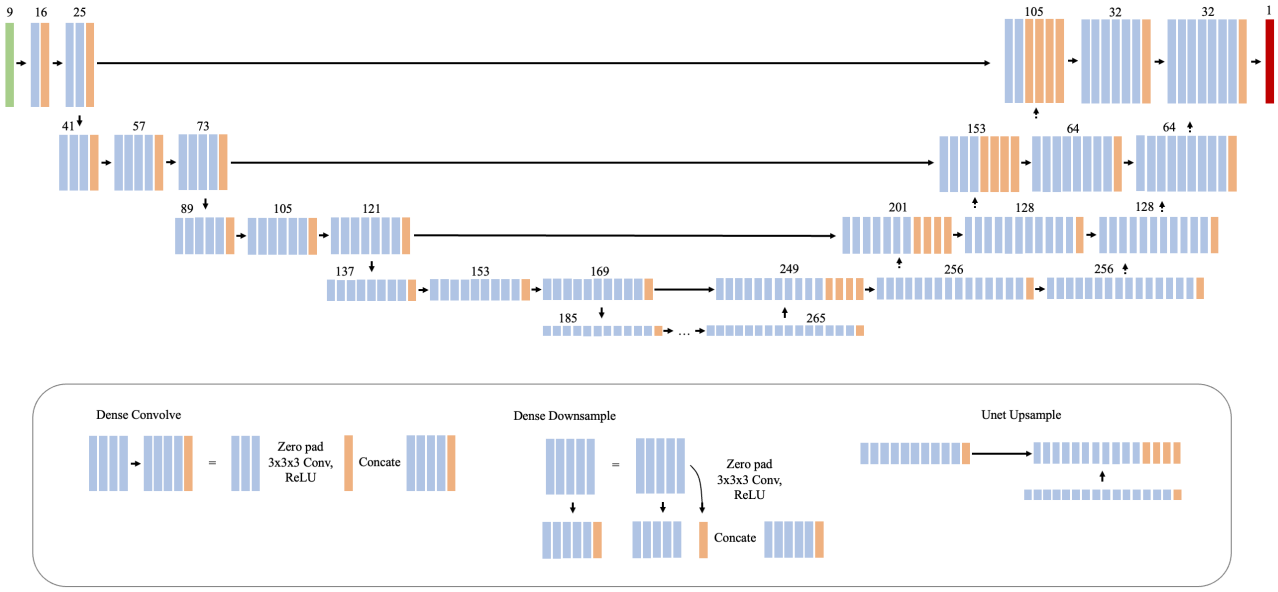


Figure 3: **Hierarchically Dense U-Net for three-dimensional dose prediction.**
Schematic of the HD-U-Net used for APBI dose prediction. Inputs are the planning CT together with binary masks for targets and organs at risk, concatenated as channels. The encoder comprises multiscale convolutional blocks with dense connections within each level and across adjacent levels to enhance feature reuse and gradient flow. Downsampling uses strided convolutions to expand the receptive field while preserving stability. The decoder upsamples features and fuses them with encoder features through skip connections and hierarchical dense links, enabling precise recovery of boundary and context information around the cavity and organs at risk.

penalties for variations. This stratification allowed us to reflect the nuanced clinical judgment typically applied in treatment evaluation.

To evaluate the distribution of plan quality across the patient cohort, we applied the scoring framework to the 550 APBI patients treated in our clinic. The scoring results showcase the inherent variance in plan quality and provide insight into how scoring metrics differentiate between acceptable and optimal plans. Key structures included PTV, heart, ipsilateral and contralateral lungs, and breast tissues.

This detailed analysis not only established the baseline quality of treatment plans but also provided a reference for the subsequent application of the scoring framework in training and evaluating AI dose prediction.

## Dose Prediction Model

Nguyen et al. proposed the use of a Hierarchically Densely Connected U-Net (HD U-Net) for predicting dose distributions in head and neck cancer patients (Nguyen et al. 2019). This architecture combines U-Net's global and local feature extraction capabilities with DenseNet's efficient feature reuse. Subsequent work extended this approach to other anatomical sites, including lung (Barrag´ an-Montero, Nguyen et al. 2018) and prostate cancer (Nguyen, McBeth et al. 2019), demonstrating the versatility of deep learning models in radiotherapy.

To predict the 3D dose distribution for accelerated partial breast irradiation (APBI) patients, we developed a deep learning model based on HD U-Net. The model incorporates dense connectivity and hierarchical upsampling, enabling efficient feature propagtion and accurate dose predictions tailored to patient-specific anatomy and clinical constraints.

The input to the model consists of concatenated CT images and structure masks, capturing the spatial relationships between the planning target volume (PTV) and surrounding organs-at-risk (OARs). The output is a 3D voxel-wise dose distribution, providing the granularity necessary for evaluating dose coverage and sparing critical structures.

The model was trained using an L1 loss function to minimize the absolute error between the predicted and clinically delivered dose distributions. We utilized the Adam optimizer with a learning rate of 0.0002 and implemented voxel-level data augmentations to enhance model robustness and generalizability. Training was conducted over multiple epochs with a dynamic batching strategy, ensuring efficient utilization of GPU resources while maintaining data diversity across iterations.

Validation was performed periodically during training to monitor generalization and prevent overfitting. The validation process included evaluating dose prediction accuracy

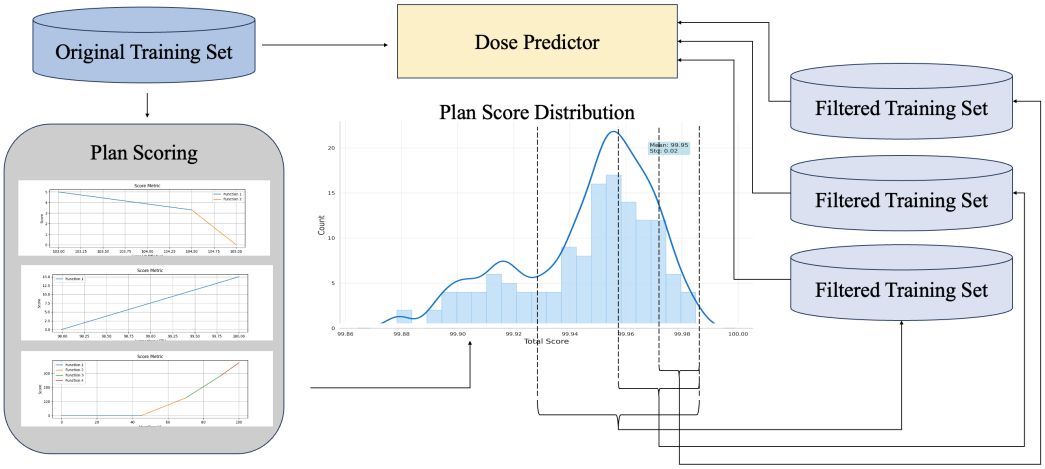


Figure 4: **Iterative refinement workflow.**
CT simulation acquires the reference anatomy, treatment planning converts clinical intent into a deliverable plan, quality assurance and plan approval confirm safety and accuracy, and treatment delivery executes the plan with image guidance and motion management.

across multiple anatomical regions, with a focus on both high-dose and low-dose regions. A checkpointing system was employed to save the model state whenever an improvement in validation loss was observed.

Post-training, the model's performance was evaluated on an independent test set. Mean squared error (MSE) and visual analyses were used to compare predicted dose distributions to clinical ground truth, ensuring both quantitative and qualitative alignment. This baseline model serves as the foundation for iterative refinement, where the adaptive scoring framework is applied to improve dose prediction performance over time.

### Iterative Refinement

We implemented an iterative quality gating procedure that progressively concentrates the training data on high scoring exemplars. An initial model was trained on the full training cohort and each case was scored with the composite plan quality metric. For the next iteration, the training set was restricted to the top 50% by score and the model was retrained. At each subsequent iteration, cases were re scored and the top 50% were again retained, yielding a progressively higher quality subset. The validation and test sets remained fixed across iterations to provide an unbiased estimate of generalization. This strategy was intended to bias learning toward superior dose patterns without losing the ability to generalize to patient plans that have a lower maximum score due to anatomical complexity.

## Results

### Patient Cohort and Scoring Distribution

Figure 5 depicts the distribution of scores across key DVH metrics for various anatomical structures, including the PTV, heart, ipsilateral and contralateral lungs, and breast tissues. The spread of scores for each metric reflects the variability inherent in clinical treatment planning due to differences in patient anatomy and the complexity of achieving optimal dose distributions.

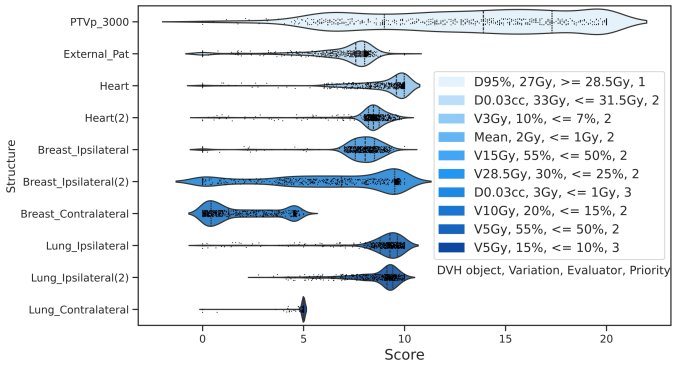


Figure 5: **Distribution of plan scores across key dose-volume histogram (DVH) metrics for the patient cohort.** Metrics include coverage of planning target volume (PTVp_3000) and sparing of OARs (e.g. heart, lungs, breast tissues). The variation in scores reflects differences in anatomical complexity and plan quality, with priority levels and acceptable threshold indicated for each metric.

Priority 1 metrics, such as D95% for the PTV, showed relatively high and consistent scores across the cohort, reflecting their critical importance in meeting clinical standards. Conversely, metrics with lower priorities, such as dose constraints for non-critical structures, exhibited broader score distributions, indicating greater variability in achieving these goals. For example, the heart and lung metrics demonstrated wider variability, particularly for metrics with lower priority, such as V5Gy and D0.03cc.

### Distribution Shifts

The performance of the dose prediction model was evaluated by comparing the predicted dose distributions against the clinical ground truth. Figure 6 presents the results from the baseline model and the following iterations.

| | mean | std | median |
|---|---|---|---|
| baseline | 43.57 | 33.10 | 57.71 |
| 1st Iteration | 57.36 | 21.77 | 64.42 |
| 2nd Iteration | 58.31 | 21.69 | 66.16 |
| 3rd Iteration | 69.39 | 11.73 | 73.72 |

**Table 2: Score statistics for fixed 88 patient test set.**

The histograms illustrate the distributions of plan scores achieved by the baseline model and the iteratively refined models. Baseline model demonstrates a broader distribution of scores, reflecting its tendency to predict dose distributions that align closely with the average quality of the clinical dataset. The iterative refined models, however, exhibits sharper distribution, indicating shifts toward higher quality predictions, as measured by the scoring framework.

The reduced spread in the score distribution for the iterative model highlights its ability to generate more consistent dose predictions, even across patients with varied anatomical complexities. Increasing alignments between predicted dose distributions and clinically desired outcomes underscores the potential for iterative refinement to improve model performance while adhering to clinical constraints.

Retraining the dose predictor with dynamically refined datasets improves the overall quality of dose predictions. The iterative training process not only elevates the predicted dose quality but also narrows the variability, ensuring more consistent performance across diverse patient cases. Future iterations aim to further optimize these gains by incorporating additional metrics addressing anatomical complexities.

## Discussion

The adaptive scoring framework presented in this study not only enhances the quality of AI-driven dose prediction but also serves as a foundation for more advanced AI methodologies in radiation oncology. By translating clinical scorecards into quantitative metrics, this framework bridges the gap between human expertise and machine learning, enabling iterative improvements in treatment planning. One of the key strengths of this approach is its potential application in reinforcement learning (RL) and continual learning (CL). In RL, the scoring framework can function as a reward signal, guiding an AI agent to explore and optimize treatment planning strategies autonomously. Similarly, CL approaches can leverage the dynamic nature of the scoring function to continuously refine AI models as new patient data becomes available, ensuring that the models remain adaptable to evolving clinical standards and patient-specific challenges.

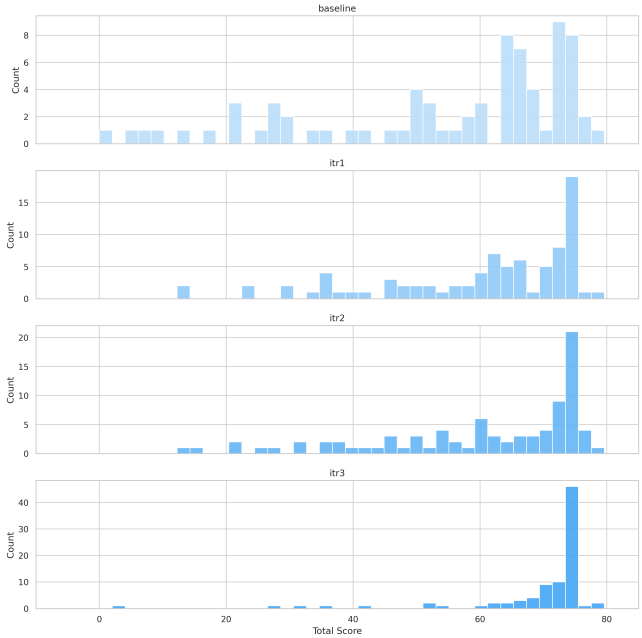


Figure 6: **Evolution of predicted plan quality across iterative training with score based curation.** Distributions of predicted plan quality scores for four models: a baseline model trained on the full training set and three successive iterations trained on the top fifty percent of cases ranked by the composite quality score from the prior round. All predictions are evaluated on the fixed test set. The plots show a progressive shift toward higher scores with reduced variability, reflected by increasing medians and shrinking interquartile ranges.

Moreover, this framework is not limited to dose prediction. It can be extended to other aspects of treatment planning, such as beam arrangement optimization, contour quality assessment, and even workflow efficiency improvements. By providing a standardized and granular method for evaluating clinical outcomes, the framework promotes greater consistency and transparency in the development of AI models, ultimately enhancing their clinical adoption.

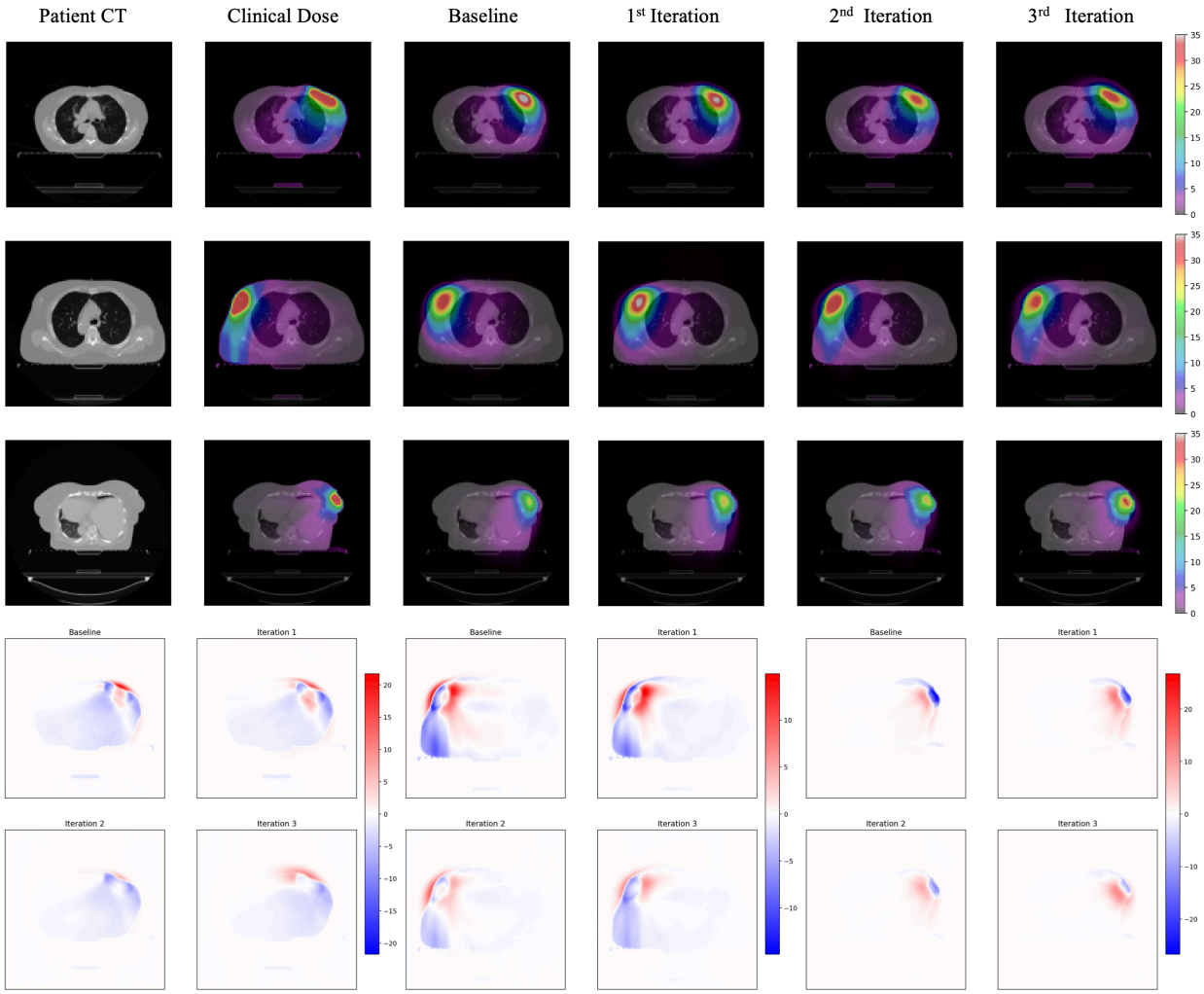


Figure 7: **Predicted and reference dose for three APBI cases with error maps.**
For three representative APBI cases, axial slices display the reference clinical dose, followed by predicted doses from four models, and the corresponding dose difference maps computed as prediction minus reference.

These advancements highlight the potential for AI to complement and augment human decision-making in radiation oncology. By aligning AI methodologies with clinical priorities, frameworks like this can drive the field toward a future where treatment planning is not only faster but also consistently achieves optimal patient outcomes.

## Conclusion

This study demonstrates the utility of an iterative refinement framework for improving radiation therapy dose prediction. By quantifying plan quality through clinically meaningful metrics, the framework enables iterative refinement of dose prediction models, leading to improved plan quality and reduced variability across diverse patient anatomies.

The results validate the hypothesis that modifying the training dataset through adaptive clinical scoring improves dose predictions quality. Potential applications of the iterative refinement extends to reinforcement learning and continual learning frameworks.

Future work will focus on several key areas:

• Expanding scoring framework to incorporate additional clinical metrics and patient-specific complexity factors.

• Exploring the integration of reinforcement learning to optimize multiple aspects of the treatment planning process simultaneously.

• Developing continual learning pipelines to adapt AI models dynamically as new clinical data emerges.

• Conducting prospective clinical trials to validate the framework's effectiveness in real-world settings.

These efforts aim to establish a robust, adaptable foundation for AI integration in radiation oncology, ultimately driving advancements in patient care through innovative technologies.

## References


Arthur, D. W.; and Vicini, F. A. 2005. Accelerated Partial Breast Irradiation as a Part of Breast Conservation Therapy. Journal of Clinical Oncology 23(8): 1726–1735. doi.org/10.1200/JCO.2005.09.045.

Barragán-Montero, A. M.; et al. 2019. Three-Dimensional Dose Prediction for Lung IMRT Patients with Deep Neural Networks: Robust Learning from Heterogeneous Beam Configurations. Medical Physics 46(8): 3679–3691. doi.org/10.1002/mp.13597.

Ge, Y.; and Wu, Q. J. 2019. Knowledge-Based Planning for Intensity-Modulated Radiation Therapy: A Review of Data-Driven Approaches. Medical Physics 46(6): 2760–2775. doi.org/10.1002/mp.13526.

Hansen, C. R.; Hussein, M.; Bernchou, U.; Zukauskaite, R.; and Thwaites, D. 2022. Plan Quality in Radiotherapy Treatment Planning: Review of the Factors and Challenges. Journal of Medical Imaging and Radiation Oncology 66(2): 267–278. doi.org/10.1111/1754-9485.13374.

Hernandez, V.; et al. 2020. What Is Plan Quality in Radiotherapy? The Importance of Evaluating Dose Metrics, Complexity, and Robustness of Treatment Plans. Radiotherapy and Oncology 153: 26–33. doi.org/10.1016/j.radonc.2020.09.038.

Kazemzadeh, A.; Rasti, R.; and Tavakoli, M. B. 2025. Artificial Intelligence for Radiotherapy Dose Prediction: A Comprehensive Review. Cancer/Radiothérapie 29(4): 104630. doi.org/10.1016/j.canrad.2025.104630.2.

Ma, J.; et al. 2021. A Feasibility Study on Deep Learning-Based Individualized 3D Dose Distribution Prediction. Medical Physics 48(8): 4438–4447. doi.org/10.1002/mp.15025.

Maniscalco, A.; et al. 2024. Multimodal Radiotherapy Dose Prediction Using a Multi-Task Deep Learning Model. Medical Physics 51(6): 3932–3949. doi.org/10.1002/mp.17115.

Nguyen, D.; et al. 2019. A Feasibility Study for Predicting Optimal Radiation Therapy Dose Distributions of Prostate Cancer Patients from Patient Anatomy Using Deep Learning. Scientific Reports 9(1): 1076. doi.org/10.1038/s41598-018-37741-x.

Nguyen, D.; et al. 2019. 3D Radiotherapy Dose Prediction on Head and Neck Cancer Patients with a Hierarchically Densely Connected U-Net Deep Learning Architecture. Physics in Medicine and Biology 64(6): 065020. doi.org/10.1088/1361-6560/ab039b.

Nguyen, D.; et al. 2021. A Comparison of Monte Carlo Dropout and Bootstrap Aggregation on the Performance and Uncertainty Estimation in Radiation Therapy Dose Prediction with Deep Learning Neural Networks. Physics in Medicine and Biology 66(5): 054002. doi.org/10.1088/1361-6560/abe04f.

Nguyen, D.; Lin, M.-H.; Sher, D.; Lu, W.; Jia, X.; and Jiang, S. 2022. Advances in Automated Treatment Planning. Seminars in Radiation Oncology 32(4): 343–350. doi.org/10.1016/j.semradonc.2022.06.004.

Njeh, C. F.; Saunders, M. W.; and Langton, C. M. 2010. Accelerated Partial Breast Irradiation: A Review of Available Techniques. Radiation Oncology 5(1): 90. doi.org/10.1186/1748-717X-5-90.

Offersen, B. V.; Overgaard, M.; Kroman, N.; and Overgaard, J. 2009. Accelerated Partial Breast Irradiation as Part of Breast Conserving Therapy of Early Breast Carcinoma: A Systematic Review. Radiotherapy and Oncology 90(1): 1–13. doi.org/10.1016/j.radonc.2008.08.005.

Xiao, Y.; et al. 2025. Embracing the Future of Clinical Trials in Radiation Therapy: An NRG Oncology CIRO Technology Retreat Whitepaper on Pioneering Technologies and AI-Driven Solutions. International Journal of Radiation Oncology, Biology, Physics 122(2): 443–457. doi.org/10.1016/j.ijrobp.2025.01.006.