Implementing Plan Robustness Evaluation in Multi-Criteria Optimization for Prostate HDR Brachytherapy

Xavier Arata[1,2], Andrew Kennedy[3,4], Luc Beaulieu[1,2]

[1] Département de Physique, de Génie Physique et d'Optique et Centre de Recherche sur le Cancer, Université Laval, Québec, Canada
[2] Département de Radio-Oncologie et Axe Oncologie du CRCHU de Québec, CHU de Québec, Université Laval, Québec, Canada
[3] School of Physics, Chemistry and Earth Sciences, College of Science, Adelaide University, Adelaide, South Australia, Australia
[4] Department of Radiation Oncology, Central Adelaide Local Health Network, Adelaide, South Australia, Australia

INTRODUCTION

- HDR brachytherapy delivers dose from within the target, giving high tumour control while sparing organs at risk (OAR).
- Every plan is subject to intrinsic, irreducible geometric and dosimetric uncertainties that can degrade its quality.
- gMCO[1] (GPU-based multi-criteria optimization) generates large pools of acceptable plans in seconds, but their robustness to these uncertainties is not currently assessed.
- **Goal:** quantify the robustness of gMCO prostate HDR plans and identify what drives it.

METHODS

- Cohort: 80 patients, US-guided HDR prostate brachytherapy (lithotomy, general anaesthesia, ~17 catheters). Contours and catheter reconstruction in Oncentra Prostate; plans from gMCO.
- The probabilistic robustness framework from Kennedy et al. is applied[2]: 500 perturbed scenarios per plan covering 7 uncertainty parameters with centre-specific data;
  - DVH recomputed via TG-43[4]; robustness = % scenarios still meeting the objectives (RTOG / INST).
- Deterministic geometric robustness from Poder et al. tested[3]: $G = \Sigma\, w / d^2$ (w = dwell weight, d = distance to the nearest prostate boundary);.
- GPU acceleration (CuPy, RTX 5090): 500 scenarios in ~1 s per plan vs ~95 s on CPU. Plan pools of 1–5000 evaluated, plus a single-uncertainty analysis isolating each parameter.

RESULTS & DISCUSSION

- Pool size (Fig. 1): robustness plateaus beyond ~50–100 plans; pools larger than ~500 add nothing.
  - ~100 plans suffice; larger pools give no benefit.

Metric	Top-robust	Top-V100	Goal
RTOG pass-rate (%)	91.5 ± 10.4	42.6 ± 8.9	—
Poder's score	3126 ± 800	3082 ± 802	—
Prostate V100 (%)	96.1 ± 1.5	98.3 ± 1.1	> 90
Prostate V150 (%)	35.8 ± 3.1	40.3 ± 3.1	—
Prostate V200 (%)	15.0 ± 2.3	16.8 ± 2.2	—
Bladder V75 (cc)	0.19 ± 0.13	0.69 ± 0.33	< 1.0
Rectum V75 (cc)	0.17 ± 0.14	0.48 ± 0.34	< 1.0
Urethra D10 (Gy)	16.5 ± 0.2	17.5 ± 0.2	< 17.7

Table 1. Dosimetric metrics (mean ± SD, 80-patient cohort): most-robust plan vs the clinically auto-selected top-V100 plan. Goal = clinical constraint.

- Table 1: the most-robust plan reaches a 91.5% RTOG pass-rate vs only 42.6% for the clinically auto-selected top-V100 plan (meeting all other dosimetric clinical criteria[1]).
- This costs only a ~2% drop in V100 (96.1 vs 98.3%, still >95%) for markedly better OAR sparing — much lower bladder/rectum V75 and urethra D10, at minimal coverage cost.
- Dose maps (Fig. 3): the robust plan spreads high dose more conservatively around the urethra and rectum.
- Single-uncertainty analysis: prostate and rectum contour uncertainties and dwell-time precision dominate pass-rate.

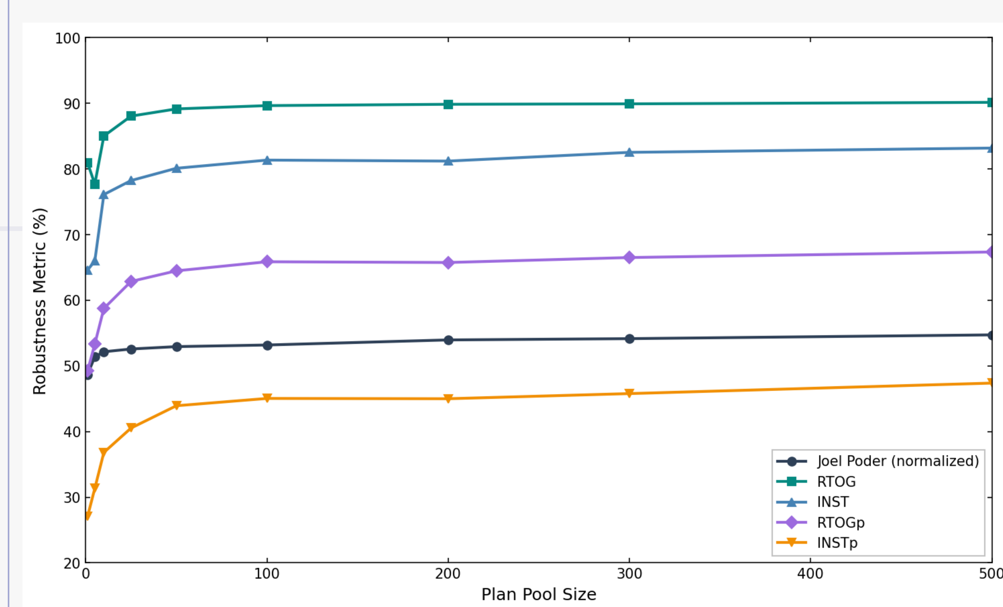

FIG. 1 - Mean maximum robustness vs plan pool size. Plateau beyond 80-100 patients

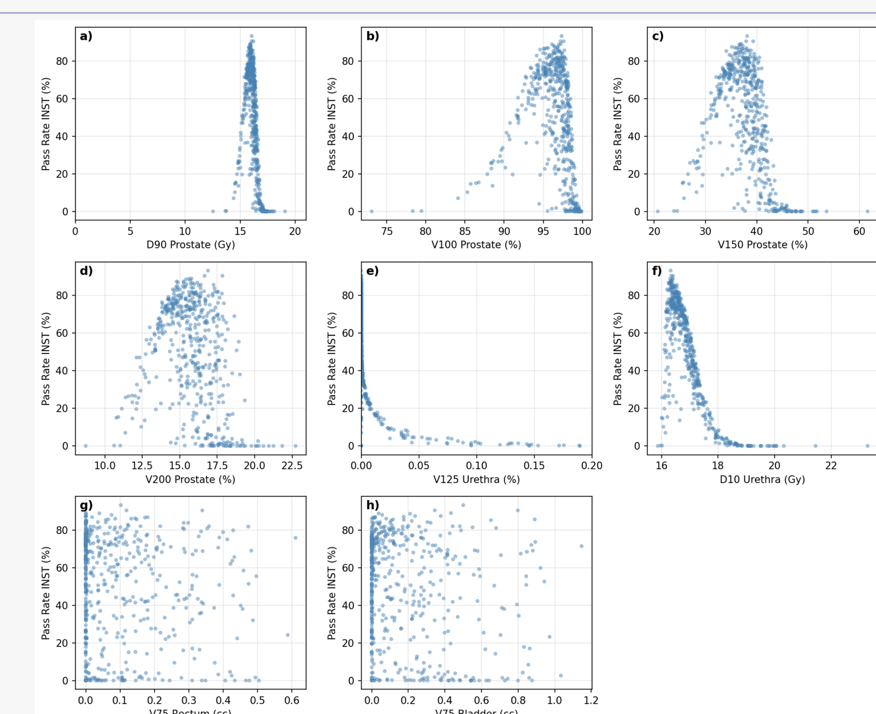

FIG. 2 – INST pass-rate vs nominal DVH metrics (1 patient, 500 plans)

Nominal DVH (Fig. 2): robustness pass-rate is driven mainly by prostate V100/V150 and urethra D10; bladder/rectum V75 correlate weakly.

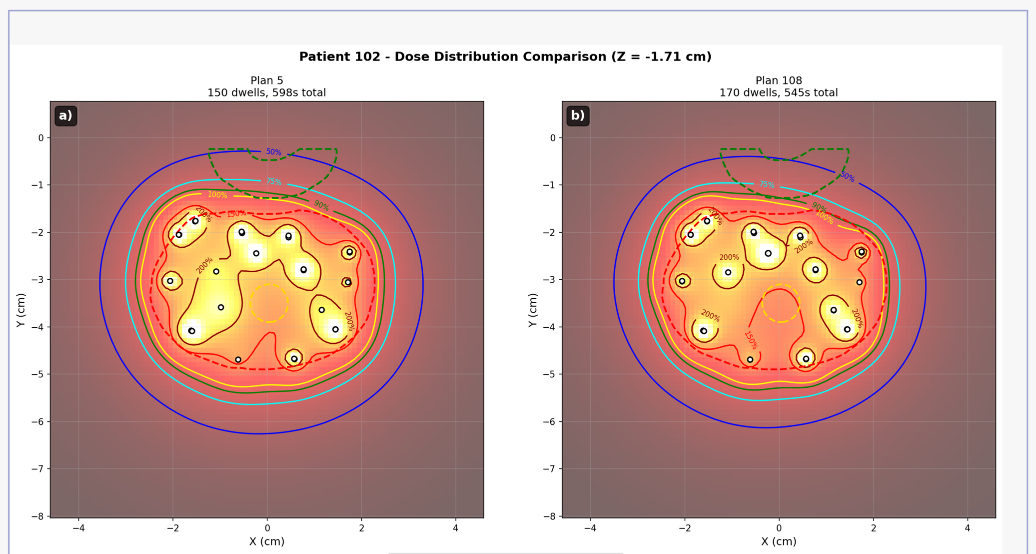

FIG. 3 – Dose distribution comparison a) most robust patient; b) top v100 patient

- Geometric vs probabilistic robustness (Fig. 4): the Poder's score and RTOG pass-rate are uncorrelated

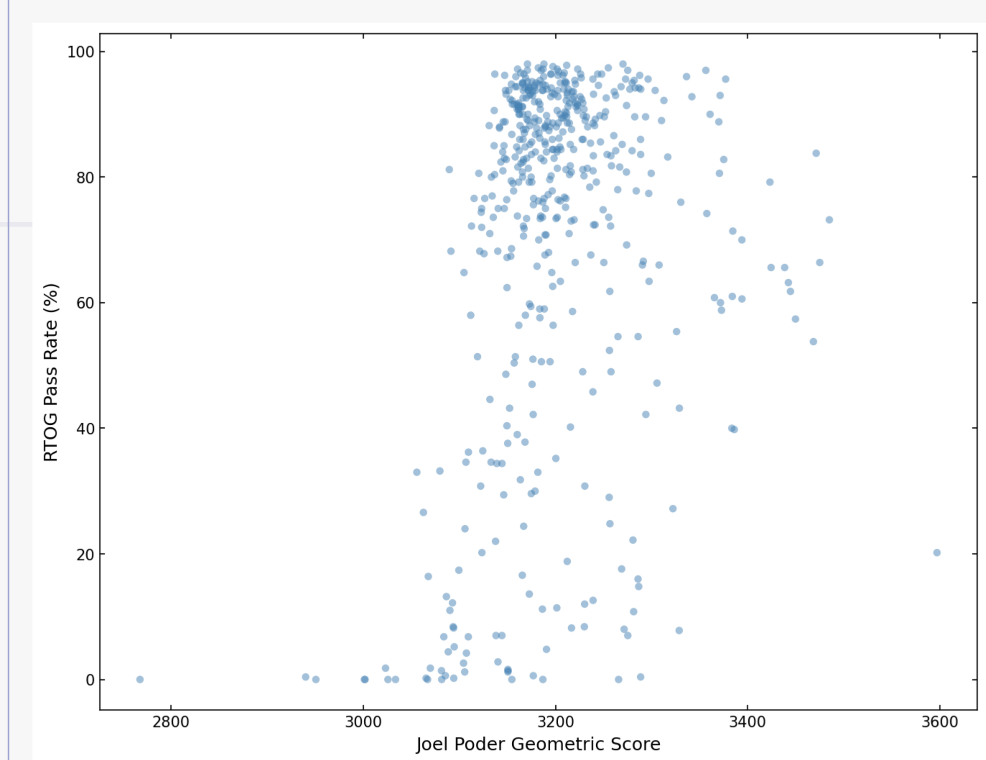

FIG. 4 – RTOG pass-rate vs geometric robustness metric

- The two metrics are not interchangeable: the geometric score uses only anatomy and dwell times, so it is fast but incomplete.
  - Refining the formula could yield a fast, computationally efficient robustness surrogate.

CONCLUSIONS

- gMCO plans can be robustly evaluated in near-real time on GPU (~1 s per plan for 500 scenarios).
- ~100 plans suffice to capture a robust plan; larger pools give no robustness benefit.
- Probabilistic (pass-rate) and geometric robustness are complementary, not interchangeable.
- The top-V100 plan is far less robust than an explicitly robust plan, at minimal coverage cost.
- Robustness feedback can be integrated into gMCO to guide robustness-aware planning.

REFERENCES

1 Bélanger et al., PMB 64 (2019) 105005.
2 Kennedy A, et al. (2024).
3 Poder J, et al. (2022).
4 Rivard MJ, et al. Med Phys. (2004).