2026 JULY 19–22 VANCOUVER, BC
JOINT AAPM | COMP MEETING

Implementation and Analysis of Deep Learning Needle Reconstruction Tool (CorneliUS) for Ultrasound-Guided HDR Prostate Brachytherapy

Yawen Shi[1]; Mathieu Goulet, PhD[2]; Robert Weersink, PhD[3,4]; Hossein Afsharpour, PhD[3,4];

[1]McGill University, Montreal, Quebec; [2]Santé Québec Chaudière-Appalaches, Lévis, Canada,
[3]Princess Margaret Cancer Centre, Toronto, Canada; [4]University of Toronto, Canada

UHN Princess Margaret Cancer Centre

Santé Québec

ABSTRACT

CorneliUS

In a previous study, CorneliUS showed time savings of 15 minutes on average in planning time for HDR prostate brachytherapy [1].

The Knowledge Gap

It is currently unknown if CorneliUS can generalize to new clinical environments without site-specific retraining.

The Method

Using an independent patient cohort from the Princess Margaret Cancer Centre, we aimed to validate the model's generalizability.

The Findings

We found similar performance on our independent dataset across needle shaft accuracy, needle detection sensitivity, and needle numbering error rates.

Take-home message

CorneliUS is highly generalizable and does not require retraining when implemented in a similar clinical setting.

CONTACT

Mathieu Goulet
Santé Québec Chaudière-Appalaches
101 Rue du Mont-Marie, Lévis,
QC G6V 0M4, Canada
Mathieu_Goulet@ssss.gouv.qc.ca
418 830-7474 ext 14539

Hossein Afsharpour
University Health Network
610 University Ave, Toronto,
ON M5G 2M9, Canada
Hossein.Afsharpour@uhn.ca
416 946-4501 ext 4273

INTRODUCTION

The Problem

Ultrasound-based HDR prostate brachytherapy requires manual needle digitization—a crucial treatment planning step that takes ~30 minutes per patient.

Study Objective

To perform the first extra-mural validation of CorneliUS, an AI-assisted digitization software that automates needles reconstruction[1], in an independent clinical setting

METHODS AND MATERIALS

Patient Cohort

Patients receiving HDR whole-gland prostate brachytherapy (15 Gy).

From Princess Margaret Cancer Centre (PMCC) : 49 patients

From Santé Québec Chaudière-Appalaches (SQCA) : 55 patients

Evaluation Metrics

AI-generated reconstructions were compared to manual (human) clinical reconstructions using four quality indicators:

MED
(Mean Euclidean Distance)
Absolute distance between AI and clinical reconstruction points (quantifies needle shaft discrepancy).

NTD
(Needle Tip Distance)
Distance (signed) between needle tip positions.

NDE
(Needle Detection Error)
Ratio of AI-detected needles to total implanted needles.

NMR
(Needle Misnumbering Ratio)
Ratio of needles correctly detected but mislabeled by the AI to total AI-detected needles

Validation & Statistical Analysis

Results from the PMCC cohort were benchmarked against baseline data from CorneliUS's original institution (SQCA).

Continuous variables (MED, NTD)
Student's t-test.

Categorical variables (NDE, NMR)
Chi-square testing.

RESULTS

Table 1. Comparison of model performance metrics between the home institution at which CorneliUS was trained on (SQCA), and our independent clinical setting (PMCC).

Performance metrics	Institution		
	PMCC	SQCA	Statistical test
MED (mm)	0.58 ± 0.47	0.59 ± 0.41	p = 0.74
NTD (mm)	3.1 ± 4.3	1.8 ± 3.9	p < 0.001
NDE	92.6%	92.7%	p = 1.00
NMR	2.8%	2.4%	p = 0.91

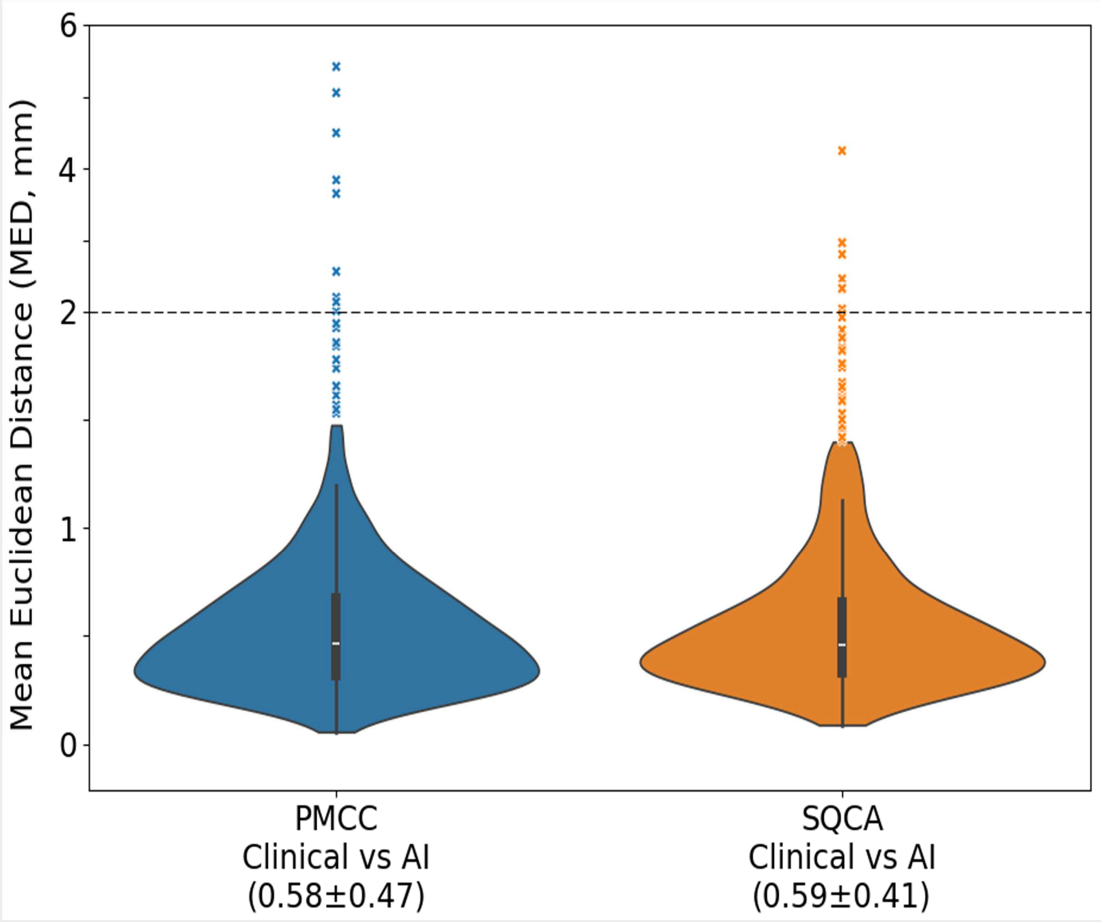

Figure 1. Resulting MED distributions for PMCC (left) and SQCA (right) for the comparison between CorneliUS-reconstructed and manually reconstructed needles. No statistical difference is found between the two institutions.

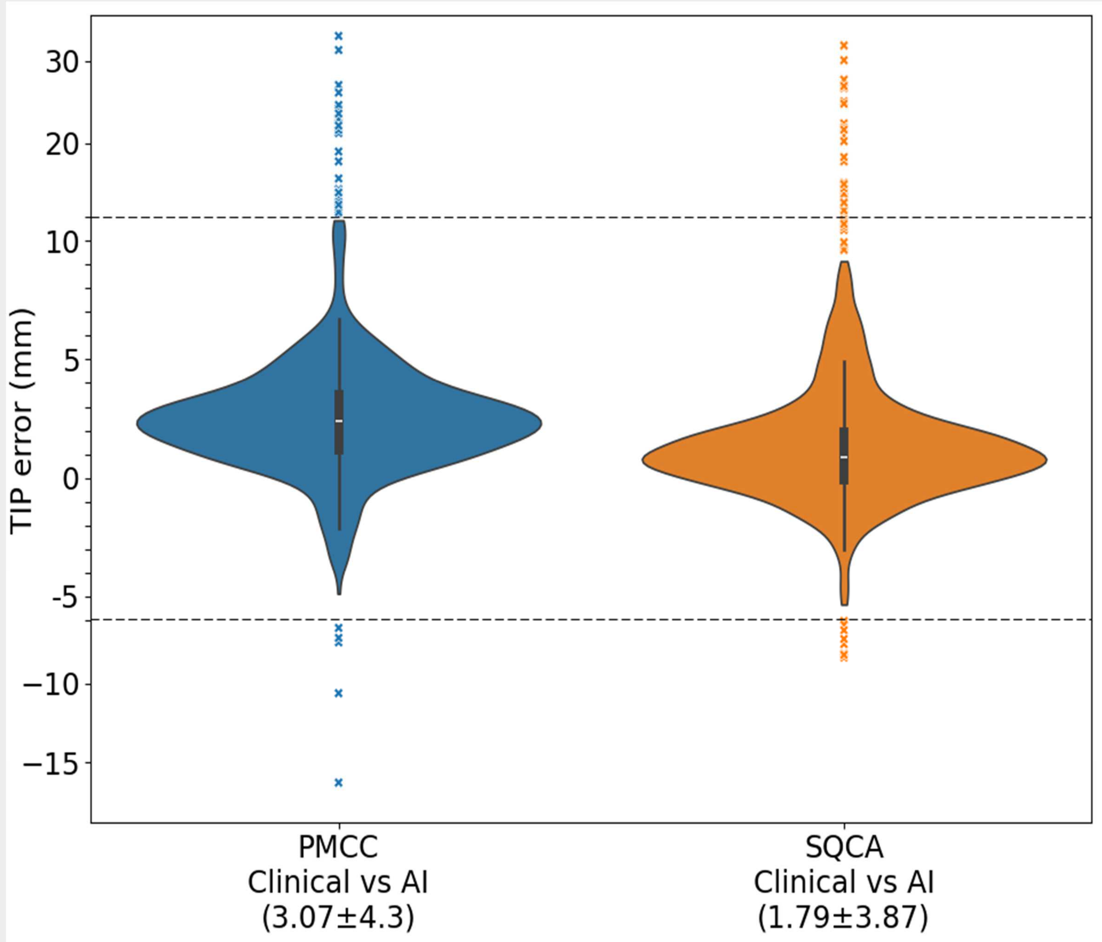

Figure 2. Resulting NTD distributions for PMCC (left) and SQCA (right) for the comparison between CorneliUS-reconstructed and manually reconstructed needles. A significant difference is found between the two institutions (p < 0.001)

DISCUSSION

Comparable Overall Performance

There was no statistical difference in Mean Euclidean Distance (MED), Needle Detection Error (NDE) or Needle Misnumbering Ratio (NMR) between the two institutions.

Needle Tip Distance (NTD)

While NTD showed a statistically significant difference between the two sites, this variation remains clinically acceptable. In standard practice, needle tips are independently adjusted and confirmed via post-implant measurements.

The Bottom Line

CorneliUS demonstrated excellent generalizability. The model achieved highly similar performance in an independent clinical setting compared to its original implementation, successfully validating its broader use.

CONCLUSIONS

Successful Validation

CorneliUS performs comparably in an independent clinical setting compared to its original implementation.

Zero Retraining Needed

The model is highly generalizable and does not require site-specific retraining when deployed in similar clinical environments.

Clinical Impact

These findings support the broader adoption of CorneliUS, highlighting its potential to significantly reduce procedure times across multiple institutions.

FINANCIAL SUPPORT

Fondation Hôtel-Dieu de Lévis

NSERC CRSNG

REFERENCES

[1] Goulet M, Duguay-Drouin P, Mascolo-Fortin J, Mégrourèche J, Octave N, Tsui JMG. Clinical Application of Deep Learning-Assisted Needles Reconstruction in Prostate Ultrasound Brachytherapy. Int J Radiat Oncol Biol Phys. 2025 May 1;122(1):199-207. doi: 10.1016/j.ijrobp.2024.12.026. Epub 2025 Jan 11. PMID: 39800329.