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

Can Enhanced-Couch Machine Performance Check Replace Off-Isocenter Winston–Lutz QA? A Data-Driven, Trigger-Based Approach for Single-Isocenter Multi-Target SRS

Ching-Ling Teng, PhD[1]; Michael D Lovelock, PhD[2]; Gary Seng Boh Lim, PhD[3]
[1]Northwell, New Hyde Park, NY 10042, [2]Mount Sina Health System, New York, NY 10029, [3]Memorial Sloan Kettering Cancer Center, New York, NY 10065

CONTACT

Name: Ching-Ling Teng, PhD, DABR
Email: cteng2@northwell.edu
Phone: +1 718 909 3689

ABSTRACT

Purpose: Off-isocenter-Winston–Lutz (OIWL) testing is increasingly recommended for single-isocenter-multi-target (SIMT) stereotactic radiosurgery (SRS), but its routine implementation is resource intensive. This pilot study evaluated whether enhanced-couch Machine Performance Check (eMPC) geometric metrics could predict OIWL quality assurance (QA) outcome and support a trigger-based QA workflow.

Methods: As a multi-institutional effort, nine paired eMPC/OIWL datasets were acquired from eight Varian TrueBeam/TrueBeam STx linacs across three health systems and six hospitals. One machine was measured before and after calibration and treated as two distinct machine states. OIWL testing was performed using the Sun Nuclear MultiMet-WL phantom. Binary OIWL QA outcome was defined as pass/fail. Associations between OIWL outcome and eMPC-reported geometric metrics were evaluated using logistic regression. Predictor screening was performed using likelihood-ratio χ2 testing and exploratory bagged-trees feature importance. For predictive modeling, ridge-regularized logistic regression was used to address complete separation. Predictive performance was evaluated using leave-one-out cross-validation (LOOCV), and operating thresholds were tuned to prioritize minimization of false-negative errors.

Results: Among all evaluated eMPC metrics, Full-Range Rotational-induced-Couch-Shift showed the strongest statistically significant association with OIWL outcome (likelihood-ratio χ2=9.53, p=0.002). In the single-variable ridge-regularized logistic model, out-of-fold predicted probabilities separated OIWL pass and fail outcomes with an AUC of 1.00 in this pilot dataset. Threshold tuning identified an operational cutoff of pOOF=0.240, corresponding to a Full-Range Rotational-Induced-Couch-Shift value of approximately 0.53 mm. At this threshold, LOOCV yielded 100% sensitivity, 100% specificity, 100% accuracy, and 100% balanced accuracy. Full OIWL required 51±17 minutes per machine, whereas eMPC required approximately 8.9±0.4 minutes.

Conclusions: In this pilot dataset, eMPC-derived Full-Range Rotational-Induced-Couch-Shift was a strong predictor of OIWL QA outcome and may support a trigger-based, risk-adaptive QA workflow in which OIWL is performed selectively when eMPC indicates elevated geometric deviation.

INTRODUCTION

SIMT SRS has been increasingly implemented across clinics for the treatment of patients with multiple brain metastases. By treating multiple lesions with a single isocenter, SIMT substantially reduces treatment time and imaging dose while maintaining conformal dose. These advantages are important for patients with advanced disease who may have limited tolerance for prolonged immobilization and are particularly beneficial for those who might otherwise receive whole-brain radiotherapy, which may adversely affect quality of life.

Despite these advantages, SIMT SRS requires submillimeter geometric accuracy not only at isocenter, but also for targets located several centimeters away. Traditional Winston-Lutz (WL) testing, however, evaluates central-axis isocenter accuracy only[1]. In SIMT SRS, geometric uncertainty increases with off-isocenter distances, and small rotational errors can result in clinically meaningful dose misalignment[2]. Therefore, in addition to AAPM TG-142 and TG-101 recommendations, AAPM-RSS Medical Physics Practice Guideline 9.b recommends additional QA for SIMT SRS[3].

OIWL testing has been developed to directly quantify targeting accuracy at multiple spatial locations. Although automated tools are available, OIWL remains resource intensive and is sensitive to phantom setup accuracy. In this study, we evaluated whether Varian enhanced-couch Machine Performance Check (eMPC) could triage the need for routine OIWL testing. We hypothesized that eMPC-reported geometric metrics could predict OIWL QA failure and support a risk-adaptive, trigger-based QA framework in which time-consuming OIWL measurements are performed selectively without compromising patient safety.

METHODS

OIWL QA was performed using the Sun Nuclear MultiMet-WL phantom and software (v2.1.0). The phantom contains six 5-mm tungsten carbide ball bearings located at isocenter up to 7 cm off-axis, enabling assessment of on- and off-isocenter targeting accuracy representative of SIMT SRS. The phantom was aligned using room lasers and CBCT. An initial acquisition was used to quantify positioning error, followed by setup adjustment. The final OIWL acquisition was analyzed after the vendor-recommended positioning accuracy criterion of ≤0.3 mm was achieved.

For each projection, a 2 × 2 cm² 6-MV MLC-defined field was delivered to each target using eight gantry/couch/collimator combinations analogous to traditional WL testing. EPID images were analyzed to determine the BB centroid, radiation-field center, field-to-target displacement, and back-projected 3D positioning and targeting errors. OIWL outcome was classified as Fail if any field-to-target displacement exceeded 1.0 mm and Pass if all field-to-target displacements were <1.0 mm. IsoCal verification was performed before each session.

eMPC was performed at the same session of OIWL QA, and ten eMPC metrics were extracted as potential predictors (Table 1). For OIWL, QA time was measured as the total test duration, including phantom setup, data acquisition, data analysis, and repeat OIWL measurements when phantom repositioning was required. For eMPC, QA time was measured as the total test duration from phantom setup to QA reporting.

Binary OIWL outcome was modeled using binomial-logit GLM and ridge-regularized logistic regression (λ= 0.001). Predictive performance was assessed using leave-one-out cross-validation (LOOCV), ROC/AUC, accuracy, sensitivity, specificity, and balanced accuracy using MATLAB R2025a.

RESULTS

- Nine paired eMPC/OIWL datasets were acquired from eight Varian TB/TB STx linacs. Sun Nuclear MultiMet OIWL QA evaluated targeting accuracy up to 7 cm from isocenter. 25% machines exceeded 1.0-mm tolerance.
- Full-Range Rotational-Induced-Couch-Shift showed the strongest association with OIWL outcome (likelihood-ratio $\chi^2 = 9.53$, $p = 0.002$), with complete separation between pass and fail groups. Beam-Center-Shift showed a weaker association ($\chi^2 = 4.24$, $p = 0.040$), whereas all other metrics were not significant (Table 1 and Fig. 1).
- Using LOOCV, the single-variable logistic model achieved an out-of-fold (LOO) AUC of 1.00. Threshold tuning identified an optimal P_{LOOF} of 0.24, corresponding to a Full-Range Rotational-Induced-Couch-Shift of 0.53 mm. At this threshold, sensitivity, specificity, accuracy, and balanced accuracy were all 100%. (Fig.2)
- eMPC required 8.9 ± 0.4 min, while OIWL test takes 51 ± 17 min. Using eMPC as a screening tool for OIWL QA, could reduce QA time by ~85%.

Predictor	Mean (SD)	Median (Range)	Single Variable	
			χ2	p-value
Full Range Rotational Induced Couch Shift (mm)	0.47 (0.23)	0.43 (0.26, 0.99)	9.53	0.002**
Beam Center Shift (mm	0.17 (0.20)	0.11 (0.05, 0.69)	4.24	0.04*
Couch Rotation (large) (deg)	0.06 (0.02)	0.06 (0.04, 0.08)	3.54	0.060
KV Imager Projection Offset (mm)	0.26 (0.06)	0.25 (0.17, 0.37)	1.83	0.176
MV Imager Projectiong Offset (mm)	0.21 (0.05)	0.2 (0.14, 0.3)	1.19	0.275
Gantry Absolute Offset (deg)	-0.06 (0.08)	-0.09 (-0.18, 0.07)	0.746	0.388
Gantry Relative Offset (deg)	-0.02 (0.10)	-0.07 (-0.16, 0.13)	0.581	0.446
Isocenter Size (mm)	0.34 (0.05)	0.35 (0.26, 0.39)	0.456	0.496
Maximum Positioning Error (mm)	0.39 (0.13)	0.35 (0.26, 0.62)	0.131	0.717
Collimator Rotation Offset (deg)	-0.01 (0.20)	-0.08 (-0.28, 0.38)	0.0316	0.859

Table 1: Single-variable analysis of eMPC geometric metrics as predictors of OIWL QA outcome. Values are presented as mean (SD) and median (range), together with the likelihood-ratio χ2 statistic and *p*-value from univariable logistic regression. Full_Range_Rotational_Induced_Couch_Shift was the strongest predictor of OIWL outcome, followed by Beam Center Shift. **p<0.05, **p<0.01.*

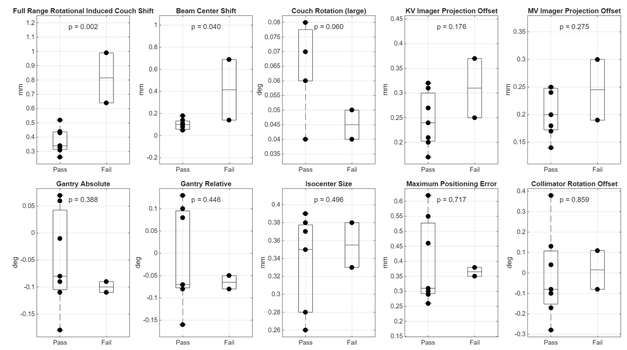

Fig 1: Distribution of Machine Performance Check (MPC) geometric metrics stratified by off-isocenter Winston–Lutz (OIWL) QA outcome (Pass vs Fail). Boxplots display median and interquartile range with individual measurements. Corresponding p-values from likelihood-ratio χ^2 testing are shown for each metric. Full-Range Rotational-Induced-Couch-Shift demonstrated the strongest association with OIWL QA outcome (p = 0.002), with clear separation between Pass and Fail groups. Beam Center Shift showed weak association (p = 0.040), while all other metrics were not significantly associated with OIWL performance.

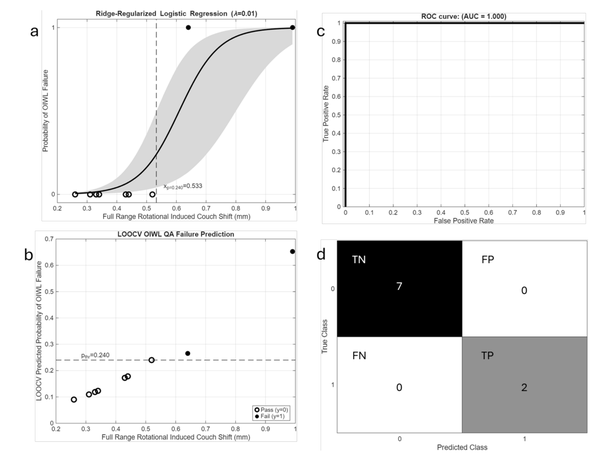

Fig 3: Predictive performance of the single-variable model using Full Range Rotational Induced Couch Shift to OIWL QA outcome. (a) Fitted ridge-regularized logistic regression curve (λ=0.01) showing the probability of OIWL failure as a function of Full Range Rotational Induced Couch Shift with 90% confidence interval. Open circles indicate OIWL Pass and filled circles indicate OIWL Fail. The dashed vertical line marks the classification threshold, corresponding to a couch-shift value of approximately 0.53 mm. (b) Leave-one-out cross-validation (LOOCV) out-of-fold predicted probabilities of OIWL failure plotted against Full Range Rotational Induced Couch Shift. The dashed horizontal line indicates the optimized operating threshold (p_{thr} = 0.240). (c) LOOCV receiver operating characteristic (ROC), demonstrating an area under the curve (AUC) of 1.000. (d) Confusion matrix at the selected threshold, showing 7 TN, 2 TP, 0 FP, and 0 FN, corresponding to 100% accuracy, 100% sensitivity, 100% specificity, and 100% balanced accuracy in this pilot dataset.

DISCUSSION

- *Can eMPC metrics predict on-axis and off-axis WL QA?* Previous studies have shown that eMPC-derived isocenter size generally agrees with on-axis WL measurements within submillimeter tolerance, but value-to-value correlations are often statistically nonsignificant[4]. Our findings were consistent with the literature. For off-axis WL, prior studies have demonstrated increasing targeting error with increasing distance from isocenter, necessitating SIMT-specific QA. Because eMPC primarily reports central-axis machine constancy metrics, it has not been considered suitable for predicting off-isocenter targeting performance. However, the present study addressed a clinically actionable endpoint: binary OIWL Pass/Fail outcome rather than continuous field-to-target displacement. In this context, Full-Range Rotational-Induced-Couch-Shift emerged as the strongest eMPC-predictor of OIWL failure. It suggests that eMPC may provide useful risk information even when direct value-to-value correlations with OIWL are limited.
- *Classification Threshold:* The proposed trigger value of 0.53 mm was selected using a risk-weighted strategy that prioritized avoidance of false-negative results. This is clinically appropriate because a false negative could permit SIMT SRS treatment despite unrecognized off-axis geometric failure, whereas a false positive primarily results in confirmatory OIWL testing. In a preliminary independent test of three additional machine states, the locked model identified the single observed OIWL failure and produced one false positive, consistent with this deliberately risk-averse strategy.
- *Risk-Adaptive Efficient QA Workflow:* Present study supports a trigger-based, risk-adaptive QA workflow in which eMPC does not replace OIWL but serves as a rapid screening tool to identify when resource-intensive OIWL testing, recalibration, or further investigation is warranted. This approach could improve efficiency while preserving safety, reducing total QA time by ~85%.
- *Limitations:* This pilot study remains limited by modest sample size, near complete separation, and limited independent testing. Therefore, the 0.53-mm trigger should be considered preliminary and requires more prospective external validation.

CONCLUSIONS

This study suggests that routinely acquired eMPC may provide a practical means of predicting OIWL QA for single-isocenter multi-target SRS. Full-Range Rotational-Induced-Couch-Shift emerged as the most statistically significant predictor of OIWL QA outcome. The addition of other eMPC metrics did not improve model performance. A clinically actionable trigger of 0.53 mm was identified, indicating that when off-axis stereotactic performance is concerned, a tighter action threshold is recommended than the vendor suggested 0.75 mm eMPC tolerance. Given its short acquisition time and reproducibility, eMPC may support a trigger-based, risk-adaptive QA workflow in which OIWL QA is performed selectively without compromising safety. Preliminary independent test supported the intended risk-averse strategy.

REFERENCES

1. Lutz W, Winston KR, Maleki N. A system for stereotactic radiosurgery with a linear accelerator. Int J Radiat Oncol Biol Phys. 1988;14(2):373-381. doi:10.1016/0360-3016(88)90446-4
2. Chang J. A statistical model for analyzing the rotational error of single isocenter for multiple targets technique. Med Phys. 2017;44(6):2115-2123. doi:10.1002/mp.12262
3. Cirino E, Benedict SH, Dupre PJ, et al. AAPM-RSS Medical Physics Practice Guideline 9.b: SRS-SBRT. J Appl Clin Med Phys. 2025;26(4):e14624. doi:10.1002/acm2.14624
4. Poder J, Chambers R, Porter H, et al. Quality assurance of single isocentre multiple target stereotactic radiosurgery: Findings from long-term off-axis Winston-Lutz testing and machine performance checks. J Appl Clin Med Phys. 2025;26(10):e70275. doi:10.1002/acm2.70275

© POSTER TEMPLATE BY GENIGRAPHICS® 1.800.790.4001 WWW.GENIGRAPHICS.COM