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

First Comprehensive Implementation of the Upcoming AAPM TG-350 Proton PBS Protocols: System Validation and Inter-room Consistency

Meiling Xu[1], Jianan Wu[1], Yajun Jia[1], Chenbin Liu[1], Yuchao Hu[1], Zhen Ding[1], Richeson Dylan[2], Peng Wang[2], Enzhuo Quan[1]*
[1]National Cancer Center/Cancer Hospital & Shenzhen Hospital, Chinese Academy of Medical Sciences & PUMC, Shenzhen, China
[2]Inova Health System, Fairfax, VA, USA

ABSTRACT

Purpose: The upcoming AAPM TG-350 provides standardized end-to-end tests for proton PBS systems. We report the first complete implementation across four treatment rooms, with emphasis on multi-room beam-matching consistency.

Methods: Seven TG-350 mock sites were planned in RayStation 7.0 and delivered to solid water using CC01 (point dose), MatrixxPT (per-field), and EBT4 film (composite). Gamma: 3%/3mm, 10% threshold. Results were compared against TG-350 benchmarks (high-dose ±4.7%, low-dose ±8.2%, per-field mean 99.0%, film mean 92.9%). One-way ANOVA tested consistency among three matched gantry rooms (GTR2/3/4) sharing an identical beam model.

Results: All rooms met or exceeded TG-350 limits. Per-field pass rates: GTR2 99.90%, GTR3 99.87%, GTR4 99.87%; film: GTR2 99.89%, GTR3 99.71%, GTR4 98.41%. ANOVA showed per-field p=0.93 (NS); film p=0.04 (significant but clinically irrelevant, all rooms far exceed TG-350 benchmark of 92.9%). FBTR1 (independent model) achieved ≥99.9% for Prostate and Pelvic sites.

Conclusion: The shared beam model delivers excellent per-field inter-room consistency (p=0.93), validating patient interchangeability among gantry rooms. Film variability reflects batch differences rather than beam delivery issues. TG-350 serves as a robust benchmark for both commissioning and multi-room QA.

CONTACT

Enzhuo Quan
quanenzhuo@cicams-sz.org.cn
National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital \& Shenzhen Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen, 518116, China CN

INTRODUCTION

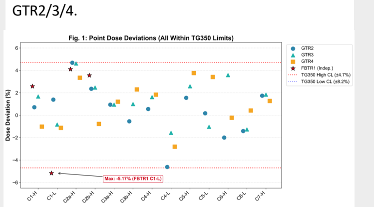
The upcoming AAPM TG-350 report provides the first standardized end-to-end testing framework for proton pencil beam scanning (PBS) systems, analogous to TG-119 for photon IMRT. It establishes multi-institutional confidence limits for point dose, per-field array, and composite film measurements across seven mock anatomical sites. However, TG-350 data were compiled from single measurements across different institutions, leaving a critical gap: **multi-room beam-matching consistency within a single center has never been systematically evaluated.**

This study presents the first complete TG-350 implementation across four treatment rooms (3 gantries + 1 fixed-beam) in a single proton center, with two primary objectives: (1) validate system performance against TG-350 benchmarks, and (2) **quantitatively assess inter-room consistency** among three gantry rooms sharing an identical beam model—addressing a clinically vital but previously unexplored question for multi-room facilities.

Our results provide the first evidence-based benchmark for multi-room PBS beam-matching, enabling confident patient interchangeability across treatment vaults.

METHODS AND MATERIALS

All tests followed TG-350 protocols on an IBA Proteus® PLUS PBS system with RayStation 7.0 TPS. **Three gantry rooms (GTR2/3/4) share an identical beam model**; a fixed-beam room (FBTR1) uses an **independent model**. Seven mock sites (C1–C7: Prostate, Pelvic Nodes, Breast, Brain, H&N, C-shape, CSI) were planned with specified SFO/MFO techniques and either Monte Carlo or analytical algorithms on a 2 mm grid (MC uncertainty ≤0.5%). Plans were delivered to a 30×30×20 cm^3 RW3 solid water phantom. **Point doses** were measured with an IBA CC01 Razor chamber (0.01 cm^3) and expressed as (Meas–Plan)/Prescription. **Per-field distributions** were acquired with MatrixxPT arrays in absolute dosimetry mode. **Composite films** used EBT4 with multi-channel calibration at 6 and 15 cm depths, scanned 24h post-irradiation. Gamma analysis employed **3%/3mm criteria** with a 10% global threshold and allowed ≤2 mm translation/rotation. Results were compared against **TG-350 published benchmarks**: high-dose ±4.7%, low-dose ±8.2%, per-field mean 99.0% (95% lower bound 95.86%), film mean 92.9% (95% lower bound 82.9%). **One-way ANOVA (α=0.05)** tested consistency among GTR2/3/4.



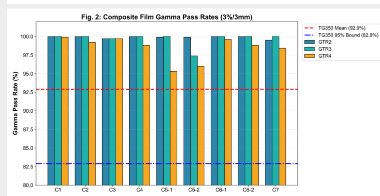
RESULTS

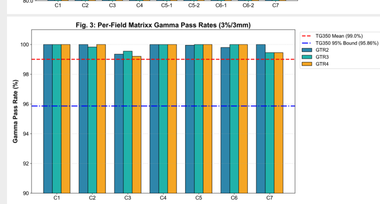
Point Dose Performance (Fig. 1)
All point-dose deviations were **within TG-350 limits**. High-dose points ranged from -2.00% to +4.68% (within ±4.7%); low-dose points from -5.17% to +3.41% (within ±8.2%). The largest deviation was FBTR1 C1-L at -5.17%. **GTR2/3/4 showed nearly identical values**, confirming excellent absolute dose transfer within the shared beam model.

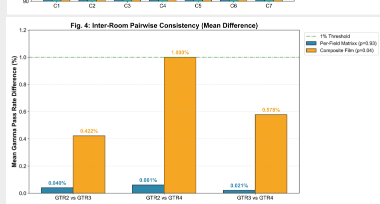
Composite Film Gamma Analysis (Fig. 2)
Mean pass rates for GTR2, GTR3, and GTR4 were 99.89%, 99.71%, and 98.41%, respectively—**all far exceeding** the TG-350 mean of 92.9% and its 95% lower bound of 82.9%. The lowest single measurement was GTR4 HN at 95.3%, still well above the benchmark. **FBTR1** (Prostate and Pelvic only) achieved **99.9% and 99.2%** , respectively.

Per-Field Matrixx Gamma Analysis (Fig. 3)
Mean pass rates for GTR2, GTR3, and GTR4 were 99.90%, 99.87%, and 99.87%, respectively—**all surpassing** the TG-350 mean of 99.0% and 95% lower bound of 95.86%. **Every individual field exceeded 98.6%. FBTR1** (Prostate and Pelvic only) achieved **100%** pass rates for all measured fields.

Inter-Room Consistency (Fig. 4)
One-way ANOVA on GTR2/3/4 gamma pass rates showed **no significant differences for per-field (p=0.93)**; film showed statistical significance (p=0.04) due to GTR4 using a different film batch with larger variability (SD=1.65%), while GTR2/3 used the same batch (SD=0.18% and 0.87%). **All rooms far exceed TG-350 benchmark of 92.9%**, confirming **excellent clinical beam-matching** among the three gantry rooms. **FBTR1** (independent model, Prostate/Pelvic only) achieved **≥99.9% pass rates** with adapted angles.







DISCUSSION

This is the **first study to systematically apply TG-350 across multiple rooms in a single center**. For the matched set (GTR2-4), inter-room ANOVA showed **no significant differences for per-field (p=0.93)**; film showed statistical significance (p=0.04), which is **attributed to GTR4 using a different film batch with larger variability (SD=1.65%)** rather than systematic beam delivery issues—critically, **all rooms still far exceed the TG-350 benchmark of 92.9%**. This provides quantitative evidence that patients can be scheduled interchangeably among these rooms without compromising plan accuracy. Point-dose measurements being nearly identical across GTR2-4 further validates the **absolute calibration transfer**—a critical requirement for multi-room facilities.

Our results **surpass TG-350 multi-center averages across all metrics**—particularly in composite film (our ~99% vs. 92.9%). We attribute this superior film performance to our **optimized multi-channel calibration protocol** with depth-matched curves (6 and 15 cm), which minimizes LET-dependent quenching effects. This methodological refinement may serve as a **useful reference for other institutions** planning TG-350 implementation.

The FBTR1, operating on an independent beam model with clinically adapted angles, also passed all applicable tests (Prostate and Pelvic) with **≥99.9% pass rates**. This suggests that **TG-350 can be reasonably adapted for fixed-beam systems**—a finding particularly relevant for centers with mixed-room configurations, though site selection requires careful consideration of beam angle constraints.

Several limitations warrant consideration. First, all measurements were performed in homogeneous phantoms, which do not fully represent tissue heterogeneity challenges. Second, FBTR1 was only evaluated for two of seven sites due to geometric constraints. Third, validation at other multi-room facilities would strengthen generalizability of our quantitative thresholds. Despite these limitations, our **inter-room consistency data establish practical benchmarks** that other proton centers can reference for validating shared beam model performance.

CONCLUSIONS

All three matched gantry rooms (GTR2/3/4) passed TG-350 benchmarks with excellent per-field consistency (ANOVA p=0.93). Although film analysis showed statistical significance (p=0.04), this reflects GTR4 using a different film batch with larger variability, not systematic beam delivery issues—and critically, all rooms still far exceed the TG-350 benchmark of 92.9%. FBTR1 (independent model) also passed all applicable tests with ≥99.9% pass rates. Overall, our system exceeds multi-center averages across all metrics, confirming that **TG-350 is a robust framework for commissioning and quantitative multi-room beam-matching QA**, with our inter-room consistency data serving as **a practical reference for other multi-room proton centers.**