FMEA severity scores: do we really know?

Jacqueline Tonigan, Jennifer Lynn Johnson, Stephen Kry, Peter Balter, Laurence Edward Court, Francesco Stingo, David Followill

Department of Radiation Physics

The University of Texas, M.D. Anderson Cancer Center, Houston, Texas

Introduction

Many physicists hold the opinion that current practices in quality management are incomplete and out dated, reflected by the 1-5% per treatment course error rates reported and the difficulty managing QM efforts for advanced techniques in busy clinics. In response, many prospective, process-wide risk mitigation techniques are becoming more prevalent in radiotherapy to establish comprehensive quality management programs, such as failure modes and effects analysis (FMEA).

Failure Modes and Effects Analysis

Potential points of failure are identified and ranked in order of the Risk Probability Number (RPN):

$$RPN = O \times D \times S$$

Where,

- O = Likelihood of occurrence
- D = Lack of detectability
- S = Severity

Scores O, D, and S are assigned by a team through consensus or average based on a scale such as that in Table 1. The most risky failure modes (high RPN) are evaluated for risk reduction.

Rank	Occurrence (O)		Detectability (D)		Severity (S)	
	Qualitative	Frequency	Qualitative	Est. probability of going undetected	Qualitative	Categorization
1	Failure Unlikely	0.01%	Never undetected	0.01%	No effect	
2		0.02%	Very low likelihood undetected	0.2%	Inconvenience	Inconvenience
3	Relatively few failures	0.05%		0.5%		
4		0.1%	Low likelihood undetected	1%	Minor dosimetric error	Suboptimal plan or treatment
5		< 0.2%		2%	Limited toxicity or tumor underdose	Wrong dose, dose distribution, location or volume
6	Occasional failures	< 0.5%		5%		
7		< 1%	Moderate likelihood undetected	10%	Recordable event, Potentially serious toxicity or tumor underdose	
8	Repeated failures	< 2%		15%		
9		<5%	High likelihood undetected	20%	Reportable event, Possible very serious toxicity or tumor underdose	Very wrong dose, dose distribution, location or volume
10	Failures inevitable	> 5%	Always undetected	> 20%	Catastrophic	

Table 1: FMEA scoring scale adopted from AAPM TG-100 and Ford, *et al.*[1]

The Issue

The subjective nature of these ordinal scores leads to variability in the scores as well as questionable reliability and validity of the results. Additionally, physics components are commonly grouped together, leaving out valuable process detail information important to physics quality management. While helpful with overall quality improvement, FMEA is lacking an objective and accurate means for analyzing current detail-oriented physics QM practices

Objective

The ***overall objective*** of this work is to reduce the subjectivity of IMRT delivery FMEA severity scores for physics components by providing quantitative data on the effects of these failures. Our ***long-term goal*** is an objective, standardized method for improvement of physics quality management. The underlying ***rationale*** is that the proposed research will allow for utilization of a new risk mitigation tool for optimization of physics QM practices.

Materials and Methods

Eleven physical failure modes (FMs) for head and neck IMRT dose calculation and delivery have been identified and are examined near commonly accepted tolerance criteria levels. Dosimetry measurements (requiring decommissioning in several cases) were performed to determine the magnitude of dose delivery errors (i.e., the severity) for six of the FMs to date. Resultant quantitative severity scores are compared to FMEA scores obtained through an international survey.

Failure Modes

1. Beam Energy
2. Beam Symmetry
3. MLC Position
4. Gantry Angle
5. Collimator Angle
6. Couch Angle
7. MU Linearity
8. MLC Transmission Modeling
9. MLC Leaf End Modeling
10. MLC Tongue-and-groove Modeling
11. CT Number-to-electron Density Table

Quantitative Assessment – Measurement

Failure modes were induced on Varian Clinac 2100CD accelerators going out of clinical service. Dosimetry measurements were made using a standard H&N IMRT plan was delivered to IROC's IMRT head and neck phantom containing TLD and radiochromic film (Figure 1). The plan was delivered as a baseline with no induced errors and then again each of the six FMs independently. Magnitude of errors induced are reported in Table 2.

Figure 1 (left): IROC's head and neck phantom

Failure Mode	Induced Error Magnitude
MLC Position	2 mm systematic, out
Beam Quality	+1.1% TMR ratio
	-0.6% TMR ratio
Symmetry	3.5%
Gantry, Collimator, Couch Angle	+2° systematic

Table 2 (above): Summary of magnitudes of induced failure modes

Qualitative Assessment – Survey

An online survey was emailed to approximately 2000 IROC-participating physicists worldwide. Physicists were asked to assign a score for occurrence, lack of detectability, and severity for each of the eleven failure modes in a IMRT head and neck case. Additionally, they assigned a percent dose error to each failure mode, indicating the worst case scenario for both PTVs and OARs.

Results

Quantitative Assessment – Measurement

Several significant differences seen between baseline and FM deliveries in TLD and film as shown below in Table 3.

Standard Treatment Plan						
Failure Mode	Average Δabs dose	Maximum Δabs dose	p	Δ DTA (mm)	p	Δ%pp
1. Beam Energy ↑	1.3%	2.2%	0.061	0.7	0.423	16%
Beam Energy ↓	1.7%	2.2%	*0.042*	0.2	0.321	9%
2. In-plane Symmetry	2.0%	3.4%	*0.015*	0.2	*0.038*	13%
Cross-plane Symmetry	3.1%	4.3%	*0.004*	0.3	0.122	18%
3. MLC Position	1.4%	2.0%	*0.005*	0. 9	*0.029*	19%
4. Gantry Angle	1.8%	3.9%	0.055	0.0	1.000	10%
5. Collimator Angle	0.3%	1.1%	0.332	0.3	0.267	0%
6. Couch Angle	0.1%	1.0%	0.377	0.0	0.742	1%

Table 3: Summary of difference between baseline and failure mode deliveries for standard H&N treatment plan. Absolute dose differences measured with 6 TLD in PTVs, p-values determined with paired t-test. DTA = Distance to agreement on axial film between primary PTV and OAR. %pp = difference from baseline in percent of pixels passing standard IROC gamma analysis with 7%/4mm criteria on axial film

Results

Qualitative Assessment – Survey

150 responses were received from North American physicists, with resultant Occurence, Detectability, and Severity scores as well as percent errors corresponding to tolerance level failures are reported below. Variability in responses was very high. Overall expected severity was low but potentially consequential, with average S ≤ 5 and average error ≤ 7%.

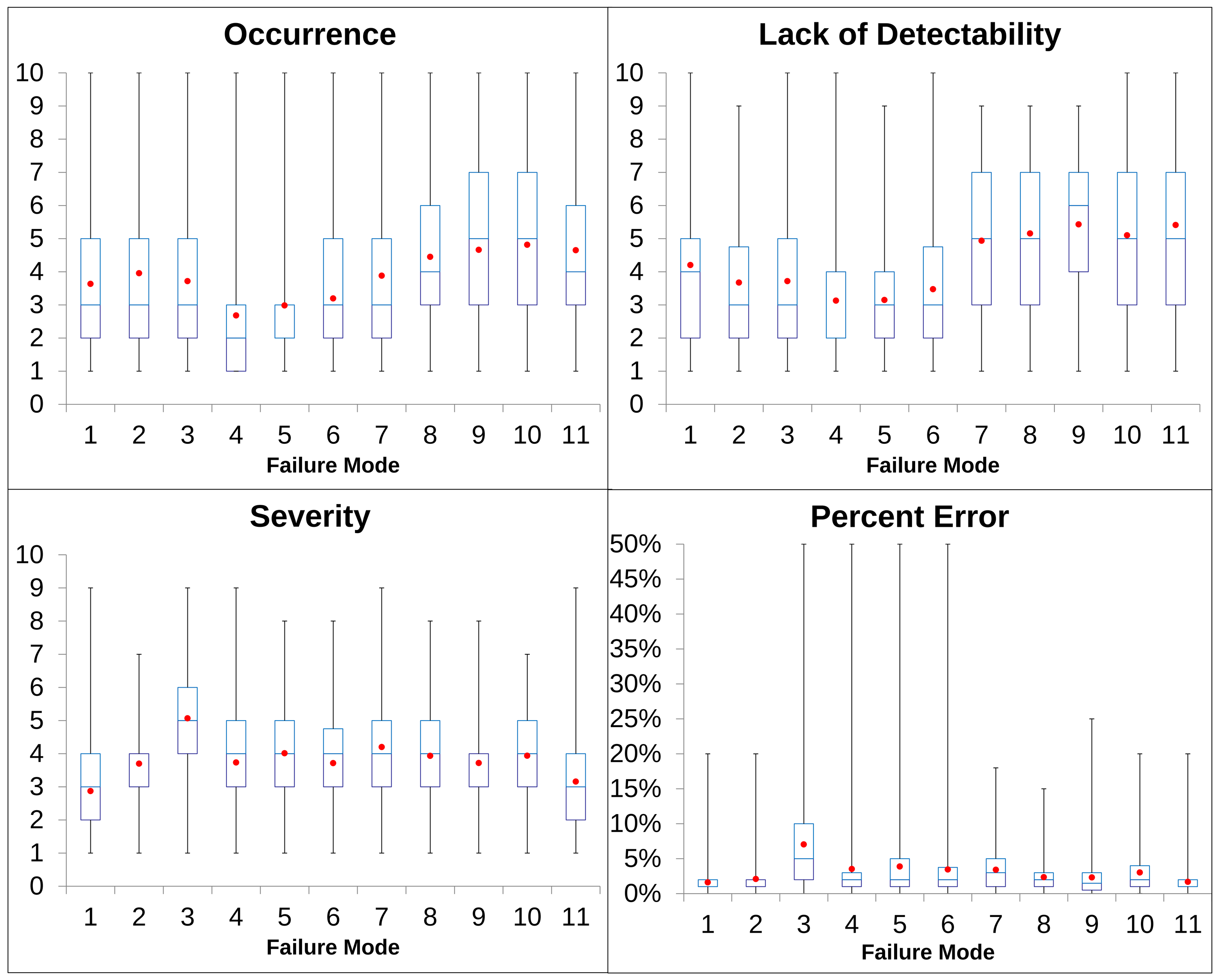

Figure 2: Whisker-Box plots of North American survey results for FMEA scores and estimated percent error for our 11 FMs. Red points show average values. N = 150

Comparison

Percent error acquired with TLD measurements are compared to those estimated in the survey to assess perception and potential validity of FMEA severity evaluation. While on the same scale, average measured severities are overall lower than estimated in survey.

FM	Quantitative (Avg. Measured Dose Error)	Qualitative (Avg. Estimated Dose Error)
1	1.3%	1.6%
2	2.6%	2.1%
3	1.4%	7.0%
4	1.8%	3.5%
5	0.3%	3.9%
6	0.1%	3.5%

Table 4: Comparison of measured and survey quantitative severity (percent dose error) for six physical failure modes near tolerance criteria levels.

Conclusion & Future Work

Significant errors have been physically measured for several of our failure modes near tolerance criteria levels, though errors are small. Large variability is noted in results of FMEA survey on these failure modes, though on average expected severity is low. Ongoing comparison of measurement, computational, and qualitative severity assessment will expand upon the functionality of FMEA in radiotherapy, allowing for the direct evaluation of current physics quality management practice and action criteria.

References

[1] Ford EC, Gaudette R, Myers L, et al. Evaluation of safety in a radiation oncology setting using failure mode and effects analysis. *Int J Radiat Oncol Biol Phys.* Jul 1 2009;74(3):

Support: *Work supported by PHS grants CA10953 and CA081647 awarded by NCI, DHHS*