Assessing Risk in Radiotherapy

Benedick A Fraass PhD, FAAPM, FASTRO, FACR
Professor, Vice Chair for Research, and Director of Medical Physics
Department of Radiation Oncology
Cedars-Sinai Medical Center, Los Angeles, CA
Clinical Health Sciences Professor, UCLA
Professor Emeritus, University of Michigan

Disclosures

• No current conflicts
• Member, TG-100
• Member, Radiation Oncology Health Advisory Council (ROHAC) for the Radiation Oncology Incident Learning System (RO-ILS) supported by ASTRO/AAPM/ClarityPSO
• Past research and/or travel funding: Varian, Elekta, Siemens, Sun Nuclear

Objectives

• Description of risk assessment methodologies used in healthcare and industry
• Discussion of radiation oncology-specific risk assessment strategies and issues
• Evaluation of risk in the context of medical imaging and image quality
Assessing Risk in Radiotherapy

- Introduction, Risk-based QM and TG-100
- Assessing Risk in Radiotherapy
- Issues for Risk Assessment and Evaluation
- Conclusions

Process-Oriented Risk-Aware Quality Methods

- Risk-aware quality methods have long been used in engineering, nuclear power plant safety, + air travel
- Goal is to improve quality and safety, while (hopefully) making our quality management efforts more efficient and effective
- This approach integrates evaluation of procedural problems with technical + device-related problems

→ TG-100

Medical Physics

After only 13 years . . .

The report of Task Group 100 of the AAPM: Application of risk analysis methods to radiation therapy quality management
Process-Oriented Risk-Aware Quality Methods

TG-100

1. Map the process to be studied
2. Analyze how the process can fail, what the effect of each failure will be, and the risk associated with those failures: (FMEA: Failure modes + effects analysis)
3. Once all the failure modes and effects are identified, map how the faults propagate: (FTA: Fault tree analysis)
4. Find efficient ways to minimize propagation of errors through the process: (QM, QA, QC)

Huq et al, TG-100, Med Phys 43: 4209 (2016)

There are lots of possible failure modes.

How do we decide what to work on first?

Let’s use Risk!

Assessing Risk in Radiotherapy

• Introduction, Risk-based QM and TG-100
• Assessing Risk in Radiotherapy
• Issues for Risk Assessment and Evaluation
• Conclusions
Assessing Risk in Radiotherapy

- Introduction, Risk-based QM and TG-100

Assessing Risk in Radiotherapy

- Scoring risk: Risk Probability Number
- What can we learn from publications?
- Incident Learning System Data
  - RO-ILS
  - Local Incident Learning Systems
- Experience

Typical Hospital Error Scoring System

<table>
<thead>
<tr>
<th>Rank</th>
<th>Qualitative Description</th>
<th>Harm</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Capacity for Error</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>Did not reach patient</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>Reached patient</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>Monitor patient</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>Intervention required</td>
<td>Temporary</td>
</tr>
<tr>
<td>6</td>
<td>Extended stay</td>
<td>Temporary</td>
</tr>
<tr>
<td>7</td>
<td>Permanent harm</td>
<td>Permanent</td>
</tr>
<tr>
<td>8</td>
<td>Intervention to sustain life</td>
<td>Harm</td>
</tr>
<tr>
<td>9</td>
<td>Death</td>
<td>Death</td>
</tr>
</tbody>
</table>

Assessing Risk in Radiotherapy

- In the context of developing our QM analysis (as in TG-100), risk will be used to prioritize issues to mitigate.
- Severity of an error is clearly important
- However, risk is more than severity, it includes the likelihood that a given error will occur
- The errors which actually cause harm are those which do not get discovered before the harm is done, so our ability to identify errors is also important
Assessing Risk in Radiotherapy

Scoring risk: Risk Probability Number

RPN = O x S x D:
- O: Frequency with which the fault occurs
- S: Severity
- D: Likelihood failure will NOT be detected

RPN = O x S x D:
- (O) Occurrence
  - 1: Failure unlikely
  - 2: Relatively few failures
  - 3: Occasional failures
  - 4: Repeated failures
  - 5: Failures inevitable

RPN = O x S x D:
- (S) Severity
  - 1: No effect
  - 2: Inconvenience
  - 3: Minor dosimetric error
  - 4: Limited toxicity or tumor underdose
  - 5: Potentially serious toxicity or tumor underdose
  - 6: Possible very serious toxicity or tumor underdose
  - 7: Catastrophic

Huq et al., TG-100, Med Phys 43: 4209 (2016)
RPN = O \times S \times D:
Detectability (D)
(Estimated probability of error going undetected)

<table>
<thead>
<tr>
<th>Rank</th>
<th>Prob. Undetected (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.01</td>
</tr>
<tr>
<td>2</td>
<td>0.2</td>
</tr>
<tr>
<td>3</td>
<td>0.5</td>
</tr>
<tr>
<td>4</td>
<td>1.0</td>
</tr>
<tr>
<td>5</td>
<td>2.0</td>
</tr>
<tr>
<td>6</td>
<td>5.0</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td>10</td>
<td>&gt;20</td>
</tr>
</tbody>
</table>

Huq et al, TG-100, Med Phys 43: 4209 (2016)

**TG-100 IMRT Analysis**

RPN Priority Score vs Rank

**Assessing Risk in Radiotherapy**

What can we learn from publications?

What do we know about treatment delivery errors?
Radiotherapy Errors
(detected with independent Record/Verify System)

<table>
<thead>
<tr>
<th>Error Rate</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>3% / Session</td>
<td>Kartha, 1977</td>
</tr>
<tr>
<td>1% / Field</td>
<td>Podmaniczky 1985</td>
</tr>
<tr>
<td>0.18% / Field</td>
<td>Macklis, 1998 *</td>
</tr>
</tbody>
</table>

* Some errors caused by R/V

The modern computer-controlled Tx delivery process has changed things

- Had opportunity to compare errors between manual and computer-controlled Tx (UM CCRS)
- All ExtBeam Txs 7/96 thru 9/97 were studied (>34k fractions)
- Tx delivery errors from QA logs, retrospective e-chart analysis, logged by therapists


Overall Error Analysis

<table>
<thead>
<tr>
<th>Errors</th>
<th>Manual Tx</th>
<th>CCRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine (%/seg)</td>
<td>M1 &gt;0.12</td>
<td>M3 &gt;0.22</td>
</tr>
<tr>
<td></td>
<td>M2 &gt;0.22</td>
<td>M4 0.03</td>
</tr>
<tr>
<td>Setup/Script (%/session)</td>
<td>&gt;0.05</td>
<td>&gt;0.17</td>
</tr>
</tbody>
</table>

While computer control has removed many of the old random delivery errors, the new process is still susceptible to systematic errors which make it thru the planning/delivery process

Process and Expectations are Important

Deviation rate as MLC technology was introduced

8/3/2016

Technology, by itself, is not the problem

<table>
<thead>
<tr>
<th>Type of Error</th>
<th>Rel. Risk (95% CL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLC</td>
<td>1.9 (1.3 - 2.9)</td>
</tr>
<tr>
<td>External Blk</td>
<td>4.4 (3.1 - 6.3)</td>
</tr>
</tbody>
</table>

• External Block required direct daily actions by RTT, while MLC was set by control system


Technology, by itself, is not the problem

<table>
<thead>
<tr>
<th>Type of Error</th>
<th>Rel. Risk (95% CL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLC</td>
<td>1.9 (1.3 - 2.9)</td>
</tr>
<tr>
<td>External Blk</td>
<td>4.4 (3.1 - 6.3)</td>
</tr>
<tr>
<td>External Wdg</td>
<td>1.3 (0.8 - 1.9)</td>
</tr>
<tr>
<td>Internal Wdg</td>
<td>2.6 (1.4 - 4.5)</td>
</tr>
</tbody>
</table>

• External Block required direct daily actions by RTT, while MLC was set by control system
• External Wdg had direct visual check by RTT, while programmed internal Wdg did not.

Errors Detected by Systematic In Vivo Dosimetry
7519 patients, in vivo dosimetry (5 years)

<table>
<thead>
<tr>
<th>Tx Preparation</th>
<th>Tx Execution</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 Prescription</td>
<td>7 Tx Setup</td>
</tr>
<tr>
<td>3 Planning</td>
<td>19 Delivery</td>
</tr>
<tr>
<td>46 Calculation</td>
<td>1 Technical Failure</td>
</tr>
</tbody>
</table>

78 / 79: involved human error

A Noel et al: Detection of errors in individual patients in radiotherapy by systematic in vivo dosimetry. Radioth + Oncol 34:144-151, 1995

How Big are the Errors?
13,385 patients, 10 years

<table>
<thead>
<tr>
<th>% of Patients Affected</th>
<th>Actual Dose Error</th>
<th>Potential Dose Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>5-10%</td>
<td>&gt;10%</td>
</tr>
<tr>
<td>0%</td>
<td>5-10%</td>
<td>&gt;10%</td>
</tr>
</tbody>
</table>

0% <5% 5-10% >10%

A big challenge:
The rate of dosimetrically-significant errors (>10%) is << 0.1 %

>>> We are looking for such errors in 1-2 patients per year in a normal clinic

Errors found in independent check

- 217 (3.5%) Any error
- 70 (1.1%) Error >5% of daily dose
- 37 (0.6%) Error >10% daily dose

3/550 patients (0.54%):
In vivo detected error after the check: two >10%


How well do independent checks work?

622 patients

Dose Error
- < 1%: 14
- 1-5%: 7
- 5-10%: 21
- >10%: 5

AS Morganti et al: Complexity index (COMIX) and not type of treatment predicts undetected errors in radiotherapy planning and delivery. Radiotherapy + Oncology 74: 283-291, 2005

Assessing Risk in Radiotherapy

Radiation Oncology Incident Learning System (RO-ILS)

- Sponsored by ASTRO and AAPM
- Managed by Clarity PSO
- Go-live: June 2014
- Participating: 104 institutions with 212 facilities
- Events reported to date: 1941
- Methods and Data analyzed by Radiation Oncology Health Advisory Board (RO-HAC)
Patient Safety Organizations (PSOs)

Made possible by Patient Safety + Quality Improvement Act (2005)
PSOs created to address needs identified in 1999 Institute of Medicine Report "To Err is Human":

• Share the goal of improving the quality and safety of health care delivery
• Collect and analyze data to identify and reduce the risks and hazards associated with patient care.
• Create a secure, non-punitive environment through confidentiality and privilege protections.

RO-ILS Quarterly Reports

• Aggregate summary of types of events, techniques, severity
• RO-HAC scores events with 5 point scale (5 is most important/severe/worrying)
• RO-HAC develops detailed descriptions and commentary for interesting or representative event reports
• Additional analysis of specific issues

RO-HAC Scoring of Events

<table>
<thead>
<tr>
<th>Severity Index</th>
<th>Severity Description</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No potential or real harm</td>
<td>Event does not pose downstream risk to world. Event is not related to patient safety or quality of care.</td>
</tr>
<tr>
<td>2</td>
<td>Mild potential or real harm</td>
<td>Event may influence the risk of other downstream events. Event may cause emotional distress or inconvenience to patient with no clinical impact.</td>
</tr>
<tr>
<td>3</td>
<td>Moderate potential or real harm</td>
<td>Event enhances the risk of other critical downstream errors. Temporary pain or discomfort for patient. Deviations from best practice, but with no obvious clinical impact.</td>
</tr>
<tr>
<td>4</td>
<td>Severe potential or real harm</td>
<td>Event with potential clinical impact that is non-critical. Limited barriers to prevention of problem.</td>
</tr>
<tr>
<td>5</td>
<td>Critical potential or real harm</td>
<td>Event with potentially critical clinical impact. Extremes related barriers to prevention of problem.</td>
</tr>
</tbody>
</table>
To evaluate what kinds of software issues were potentially involved in the various events in the RO-ILS database:

- Studied 167 events with highest scores (4, 5)
- For each event, assign category, then guesstimate one or two software issues which appear to be involved
- Result is a matrix of categories vs issues, with # of events and their score
There is much more useful information about the weaknesses in our processes and our software systems, as long as
• We get detailed reports about what actually went on for each event, and
• We are clever enough to aggregate the specific problems into reasonably well-defined areas that need attention within the software designs and implementations.

Assessing Risk in Radiotherapy
Incident Learning System Data: Local Systems

In the end, it’s what we do locally that counts

Setup+Prescription Errors (%/session)

<table>
<thead>
<tr>
<th>Manual Tx</th>
<th>CCRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Setup + Script Errors</td>
<td>M1</td>
</tr>
<tr>
<td>Patient setup</td>
<td>.03</td>
</tr>
<tr>
<td>Patient/Plan choice</td>
<td>0</td>
</tr>
<tr>
<td>Prescription/Chart</td>
<td>.01</td>
</tr>
<tr>
<td>Total/session (%)</td>
<td>.05</td>
</tr>
</tbody>
</table>

1 specific process problem: 90% of these errors

Fraass IJROBP 42 (1998)
Setup+Prescription Errors (%/session)

<table>
<thead>
<tr>
<th></th>
<th>Manual Tx</th>
<th>CCRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Setup + Script Errors</td>
<td>M1</td>
<td>M2</td>
</tr>
<tr>
<td>Patient setup</td>
<td>.03</td>
<td>.07</td>
</tr>
<tr>
<td>Patient/Plan choice</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Prescription/Chart</td>
<td>.01</td>
<td>.10</td>
</tr>
<tr>
<td>Total/session (%)</td>
<td>.05</td>
<td>.17</td>
</tr>
</tbody>
</table>

1 specific process problem: 90% of these errors
Fixing this one process problem changes these results significantly

Assessing Risk in Radiotherapy

- Introduction, Risk-based QM and TG-100
- Assessing Risk in Radiotherapy
- Issues for Risk Assessment and Evaluation
- Conclusions

Issues for Risk Assessment and Evaluation

Analyzing risk: some results from the TG-100 Example FMEA of IMRT

<table>
<thead>
<tr>
<th>Rank</th>
<th># RPN</th>
<th># RPN</th>
<th># FMEA</th>
<th># Sub Process</th>
<th>Step</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>388</td>
<td>31</td>
<td>F1. Other Pre-TRT Imaging for CTV Localization</td>
<td>FM: Incorrect interpretation of tumor or normal tissue</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>10</td>
<td>F2. CT-im</td>
<td>Image not saved or sent to Tx Planning</td>
<td>FM: Simulation data accidentally deleted</td>
</tr>
</tbody>
</table>

Fraass IJROBP 42 (1998)
How Should We Prioritize Issues to Fix?

TG-100: Priorities for Mitigation

There are many valid ways to prioritize problems:
1. RPN (Risk Priority #) = O x S x D
2. Severity
3. Other

- TG-100 used RPN and then added in high severity issues (even if their RPN's were low).
- One of the most difficult prioritization issues: those FMs which are catastrophic but highly unlikely (high S, low O and D)

How Should We Prioritize Issues to Fix?

Catastrophic Errors and ‘Risk’

What do we do with catastrophic or “Never” events?

For true “never” events: Is it enough to just rank this “risk” very high, even though the occurrence or frequency is expected to be extremely small?

Or should we:
- Do rigorous hazard analysis
- Change the process to avoid potential triggers
- Develop/implement rigid QC process immediately before potential triggers

Let’s discuss a few high priority example failure modes, and understand what led those problems to be scored as “high risk”
### The top 100 Failure Modes

<table>
<thead>
<tr>
<th>Rank</th>
<th>Failure Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>#1 FM: Pre-Tx imaging: Incorrect interpretation of tumor or normal tissues.</td>
</tr>
<tr>
<td>2</td>
<td>#2 FM: GTV/CTV: Large segmentation errors: wrong organ, wrong site, wrong expansions.</td>
</tr>
<tr>
<td>3</td>
<td>#3 FM: Tx Planning Directive: Wrong or missed summary of previous radiation treatment.</td>
</tr>
<tr>
<td>4</td>
<td>#4 FM: GTV/CTV: Excessive delineation errors in target contouring.</td>
</tr>
<tr>
<td>5</td>
<td>#5 FM: CT/CTV: Other delineation errors in target contouring.</td>
</tr>
</tbody>
</table>

### The common threads for many of these (and other) high priority failure modes are:

- Problems that lead to systematic error in Tx for the whole treatment course
- Many of the issues are geometric, not dosimetric
- Decisions that are qualitative and/or dependent on complex clinical judgment (i.e., not technical)
- Failure early in the planning process is unlikely to be detected by anyone later
- Technical QA checks are unlikely to identify many of these problems

### Qualitative issues are a majority of the top 100 failure modes!

![Graph showing qualitative issues]

- % MachineQA
- % Decisions

<table>
<thead>
<tr>
<th>Failure Modes related to Decision(s)</th>
<th>% MachineQA</th>
<th>% Decisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>25%</td>
<td>25%</td>
<td>25%</td>
</tr>
<tr>
<td>50%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>75%</td>
<td>75%</td>
<td>75%</td>
</tr>
<tr>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

**IMRT analysis**
One reason: Bad decisions are hard to detect

Lack of process steps which can detect poor or incorrect decisions is a significant component of our risk

issues for risk assessment and evaluation

Balancing Risks

- Our risk environment is different from typical industrial applications of FMEA: they have normal people as their "target".
- In the RadOnc environment, we must balance risks – the risk of problems during treatment compared to the clear risk of major problems if treatment does not happen. This is a very different tradeoff situation.
- It is important to proceed with care as we implement risk-based QM so we accommodate the (different) risk environment that we operate in.

issues for risk assessment and evaluation

Risks We Can Affect

- This FM is very large target contouring errors (>3x the expected inter-operator delineation error)
- However, definition (as opposed to delineation) is one of the weakest links in our whole process, something that needs to be addressed with new science. We cannot truly resolve this fundamental problem until we know the right answer!
Assessing Risk in Radiotherapy

- Introduction, Risk-based QM and TG-100
- Assessing Risk in Radiotherapy
- Issues for Risk Assessment and Evaluation
- Conclusions

Conclusions

- TG-100 introduces medical physics to a risk-based paradigm for design of QM programs
- Here, we discuss “risk” as a metric used to prioritize issues to mitigate
- Risk is a combination of many factors, including (at least) severity, frequency, and our ability to detect a given error when it occurs
- Risk is not absolute: risk associated with treatment for cancer has to be traded off against the (usually very bad) outcome when there is no treatment.
**Tx Delivery Error Analysis**

Machine Errors (%/Segment)

<table>
<thead>
<tr>
<th>Machine Errors</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual Tx</td>
<td>.02</td>
<td>.02</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>CCRS</td>
<td>.02</td>
<td>.02</td>
<td>.03</td>
<td>.02</td>
</tr>
</tbody>
</table>

Machine Setup

<table>
<thead>
<tr>
<th>Accessory</th>
<th>Manual Tx</th>
<th>CCRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient setup</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>Patient/Plan choice</td>
<td>.09</td>
<td>.02</td>
</tr>
<tr>
<td>Prescription/Chart</td>
<td>.09</td>
<td>.02</td>
</tr>
<tr>
<td>Total/Segment (%)</td>
<td>.12</td>
<td>.22</td>
</tr>
</tbody>
</table>

Expect that these errors are under-reported, probably are 1-2%.

---

**Setup+Prescription Errors (%/session)**

Almost no way to find errors except weekly portal images.

<table>
<thead>
<tr>
<th>Setup + Script Errors</th>
<th>Manual Tx</th>
<th>CCRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient setup</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>Patient/Plan choice</td>
<td>.09</td>
<td>.02</td>
</tr>
<tr>
<td>Prescription/Chart</td>
<td>.09</td>
<td>.02</td>
</tr>
<tr>
<td>Total/Session (%)</td>
<td>.05</td>
<td>.17</td>
</tr>
</tbody>
</table>

No way to identify these manual errors.

- Includes any table coord issue (1 mm)
- 1 specific process problem: 90% of these errors

---

**Event Category**

Sum of event scores in each matrix element
There are many paths thru the IGRT process, and many places that something unexpected happens.

Inconsistencies involving the prescription, plan, scheduling, and/or documentation is a major problem.

Verbal instructions, poor communication, lack of updating all systems, changes not implemented, and the inability of many systems to follow through all updates for a change (including notifying all the right people) continues to be a big issue.
• Detailed description of 216 failure modes (FMs), their potential causes and effects, and examples
• Ranking of FMs
• Example parts of a QM program derived (in part) from the analysis
• Discussion and analysis of the RPN estimations performed by the task group
• Discussion of trade-offs involved in the different methods of prioritization of mitigation efforts
• Broad analysis of many issues highlighted by the example FMEA, FTA, + QM program development