Risks and Benefits of Automation in Radiation Therapy

James M. Lamb, Ph.D
Department of Radiation Oncology
University of California, Los Angeles
AAPM 61st Annual Meeting
July 16th, 2019

Disclosures

• Research funding from ViewRay, Inc.
• Consulting fees from ViewRay, Inc.
• AHRQ R01 HS026486

Automation: Using technology to perform tasks with minimal human effort.
Gross Injurious Machine Malfunction

- Between 1985-1987, 6 people killed by massive overdoses from Therac-25 therapy linacs

  • Contributing causes:
    - Bad design and poor testing procedures by the manufacturer
    - Inadequate reporting system.

Nancy Leveson, University of Washington

Data transfer between systems

Incorrect procedures / inappropriate training
Conventional Fumbles

**Reason's Swiss Cheese Model**

The Swiss cheese model of how defenses, barriers, and safeguards may be penetrated by an accident trajectory.


Hierarchy of effectiveness
We are developing an algorithm to automatically detect patient ID and serious patient positioning errors by analyzing image similarity of setup images with planning CTs.

We will use our algorithm to make an on-line never-event prevention system (NEPS) that will interlock the treatment machine until the right patient is in the right treatment position.
AHRQ R01 Project

- **Specific Aim 1:** We will develop a robust never-event detection algorithm for planar x-ray setup imaging.

- **Specific Aim 2:** We will employ our never-event detection algorithms to measure the clinical never-event rate [in UCLA and VA image databases], testing the hypothesis that it has been significantly underreported.

- **Specific Aim 3:** We will implement our never-event detection algorithms in an on-line, real-time never-event prevention system (NEPS) at UCLA to test its feasibility in a clinical environment.

---

**Preliminary Results (Aim 1)**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Accuracy</td>
<td>97.60%</td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>97.23%</td>
<td>97.26%</td>
</tr>
<tr>
<td>Specificity</td>
<td>97.66%</td>
<td>96.31%</td>
</tr>
</tbody>
</table>
Automation in Clinical Radiotherapy Workflow

Segmentation
Planning
Plan Evaluation
Weekly chart checks
Machine QA

UCLA Experience With Clinical Automated Planning

• Clinically validated a publically available RapidPlan prostate model, including physician determination that results were clinically acceptable.

• 1 dosimetrist likes to use it, 2 dosimetrists tried it and don’t want to use it, 4 dosimetrists didn’t want to try it.

• Currently implementing RapidPlan model trained on UCLA data

UCLA-Specific RapidPlan Prostate SBRT Model

• Trained on 50 cases from one physician, planned by one dosimetrist

• 17 model iterations to optimize DVH matching to 10 validation cases.

• 10 test cases used for “Auto-Planning Turing Test”
UCLA-Specific RapidPlan Prostate SBRT Model

When applied to outliers in historical record, improves on clinical plan.

RapidPlan significantly lower on rectum

UCLA Experience with Clinical Contouring Automation

"MIM Assistant" Adds auto-contours, keeping only every Nth slice — key development

Physician

Dosimetrist

Showing auto-contours and final approved contours

Only used for a subset of contours that work reasonably ok.

Only keep every Nth (2nd/3rd/7th) slice to make corrections easier.
Human factors barriers to RT planning automation
An interview with UCLA medical dosimetrists

- Inability to ascertain whether it helps/hurts time
- Perfectionism
- Sensitivity to criticism / power difference
- Fear of liability
- Fear of losing skills
- Fear of losing intellectual work product

Human factors barriers to RT planning automation
An interview with UCLA medical dosimetrists

- Perfectionism
- Sensitivity to criticism / power difference
- Fear of liability
- Fear of losing skills
- Fear of losing intellectual work product

“...Yes I know 1-2 mm doesn’t matter because all the organs are going to be a bit different at treatment time. But I’m supposed to make my part as precise as possible.”
Human factors barriers to RT planning automation
An interview with UCLA medical dosimetrists

- Perfectionism
- Sensitivity to criticism / power difference
  “If the contours don’t look right, even if it doesn’t matter, the doctor will think I don’t know how to do my job.”
- Fear of liability
- Fear of losing skills
- Fear of losing intellectual work product

“EZfluence works really well for simple breast cases, but I worry that if I get used to it, I’ll get out of practice, and it will make it harder to do complicated breast cases.”
- Fear of losing intellectual work product
Human factors barriers to RT planning automation
An interview with UCLA medical dosimetrists

- Perfectionism
- Sensitivity to criticism / power difference
- Fear of liability
- Fear of losing skills
- Fear of losing intellectual work product

“If I could imagine working in a place where everybody just pressed a button to optimize plans, I would hate working in that place.”

Although intended to reduce errors, automation can also induce errors.

- Alert fatigue
- Over-reliance on technology

Elimination of error checks performed by humans

Alert fatigue

- Classic examples:
  1. Beeping physiological monitors
  2. Drug-drug interaction alerts in CPOE systems

Also implicated in R&V overrides

*Alert fatigue* linked to patient’s death
98-year-old resident of a care facility

15 hours after she was placed in the facility’s hospital, the patient was pronounced dead.

Joint Commission Sentinel Event Alert
Over-reliance errors

**Tesla’s Self-Driving System Cleared in Deadly Crash**

“Tesla had said its camera failed to recognize the white truck against a bright sky. But the agency essentially found that Mr. Brown was not paying attention to the road. It determined he... should have had at least 7 seconds to notice the truck before crashing into it.” – NYTimes

**Image: NTSB**

**Safety driver of fatal self-driving Uber crash was reportedly watching Hulu at time of accident**

Over-reliance errors in RT

**FACILITATION OF RADIOTHERAPEUTIC ERROR BY COMPUTERIZED RECORD AND VERIFY SYSTEMS**

Gordet A. Patton, M.D., M.S., M.B.A.,” David K. Gettry, M.D., Ph.D.,” and John H. Mensin, M.S.

“The common denominator among these R&V-related errors was excessive reliance upon the computer system by therapists.”

“R&V cannot substitute for thinking on the part of radiation therapy team members.”

Over-Reliance Errors in RT

<table>
<thead>
<tr>
<th>Plane Quality Report</th>
<th>Pride - Princess NBT1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helmed:</td>
<td>0</td>
</tr>
<tr>
<td>Source / Film:</td>
<td>2020.02.09 - 320mm</td>
</tr>
<tr>
<td>Scale:</td>
<td>200.9%</td>
</tr>
<tr>
<td>Report Creation Date :</td>
<td>6/10/2019</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Structure</th>
<th>Custom</th>
<th>Calculated Ideal Cost</th>
<th>Acceptable Cost Achieved</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body</td>
<td>Body</td>
<td>85.10%</td>
<td>&lt;6.0%</td>
<td>None</td>
</tr>
<tr>
<td>Roccas</td>
<td>O_Rocc</td>
<td>V20.Gy</td>
<td>&lt;20%</td>
<td>None</td>
</tr>
<tr>
<td>Roccas</td>
<td>O_Rocc</td>
<td>V10.Gy</td>
<td>&lt;10%</td>
<td>None</td>
</tr>
<tr>
<td>Roccas</td>
<td>O_Rocc</td>
<td>V10.Gy</td>
<td>&lt;5%</td>
<td>None</td>
</tr>
<tr>
<td>Roccas</td>
<td>O_Rocc</td>
<td>V10.Gy</td>
<td>&lt;5%</td>
<td>None</td>
</tr>
<tr>
<td>Roccas</td>
<td>O_Rocc</td>
<td>V10.Gy</td>
<td>&lt;5%</td>
<td>None</td>
</tr>
<tr>
<td>Roccas</td>
<td>O_Rocc</td>
<td>V10.Gy</td>
<td>&lt;5%</td>
<td>None</td>
</tr>
<tr>
<td>Shoulder</td>
<td>O_ShO</td>
<td>V20.Gy</td>
<td>&lt;20%</td>
<td>None</td>
</tr>
<tr>
<td>Shoulder</td>
<td>O_ShO</td>
<td>V10.Gy</td>
<td>&lt;10%</td>
<td>None</td>
</tr>
<tr>
<td>Shoulder</td>
<td>O_ShO</td>
<td>V10.Gy</td>
<td>&lt;5%</td>
<td>None</td>
</tr>
<tr>
<td>Shoulder</td>
<td>O_ShO</td>
<td>V10.Gy</td>
<td>&lt;5%</td>
<td>None</td>
</tr>
<tr>
<td>Forearm</td>
<td>O_Fore</td>
<td>V10.Gy</td>
<td>&lt;10%</td>
<td>None</td>
</tr>
<tr>
<td>Forearm</td>
<td>O_Fore</td>
<td>V10.Gy</td>
<td>&lt;5%</td>
<td>None</td>
</tr>
</tbody>
</table>
Final thoughts

• Essential to consider human factors when implementing automation in the clinic.

• Introducing automation will prevent some errors, cause others

• In the long run, automation will be a force for good

How do we learn about errors?
• Academic studies of reported error rates of 1-4% per patient including errors with little/no clinical impact
• Per-patient reported error rates in the range of 0.1% for errors resulting in under/over-dose by 5% or more [1,2].
• It is likely that errors are under-reported [3-5]
• In-vivo dosimetry studies have shown ~0.5% per patient rate of unreported and otherwise undetected “serious” errors [6]


AHRQ R01

• Specific Aim 1: We will develop a robust never-event detection algorithm for planar x-ray setup imaging.
• Specific Aim 2: We will employ our never-event detection algorithms to measure the clinical never-event rate [in UCLA and VA image databases], testing the hypothesis that it has been significantly underreported.
• Specific Aim 3: We will implement our never-event detection algorithms in an on-line, real-time never-event prevention system (NEPS) at UCLA to test its feasibility in a clinical environment.

Another actual incident report

1. I brought patient John Doe into the room with the vacloc setup and covered with the sheet. Frank and I set up the patient and we went outside the room.
2. Frank is driving and doing the MVCT on the Tomotherapy machine, and I was doing the encounter, timeout and charting.
3. Once images were done and matched, and saved, the images were sent over. I was importing the MVCT image document and noticed the vertical shift was over 10mm tolerance (14mm) approx., and told my coworker to stop the treatment machine.
Independent, automated, failsafe backups of human decision making, while still leaving the human in control of the workflow.

- Construction management systems
- Ground proximity warning systems for avoiding "controlled flight into terrain"
- Air traffic control conflict detection systems
- Physician decision support, e.g., detecting harmful drug-drug interactions and missed allergies using computerized order entry.

An actual incident report

1. Exactrac automatic fusion resulted in a 20 – 25 mm longitudinal misalignment that was not caught by the treating therapists. The misregistration performed by automatic alignment may have been caused by the initial vertical offset. The large shift was cause for concern and shifted the patient based on the 25 mm misregistration. During follow-up verification imaging they then concentrated on aligning the two wrong vertebral bodies. When they couldn’t get a good alignment they proceeded to image and align with CBCT.

2. When CBCT was performed the patient was already misaligned by 1 vertebral body. The therapists did not observe the large discrepancy during CBCT image registration and again concentrated on a limited field of view as they attempted to fuse the two mismatched vertebral bodies.

3. An therapist unfamiliar with Exactrac imaging was filling in on the machine during the treatment. Thus, only the therapist driving the Exactrac IGRT was familiar with the process. The fill-in therapist was not certain of the alignment but defaulted to the judgment of the Exactrac therapist. Thus, there was only limited "second check" of alignment at the machine.
Reason’s psychology of errors

Some of Reason’s Recommendations (Reason 1995):
- “Human rather than technical factors represent the greatest threat to complex and potentially hazardous systems.”
- “Effective risk management depends on a confidential and preferably anonymous incident reporting system.”
- “Automation and increasingly advanced equipment do not cure human factors problems.”

Errors (as defined above)

Vs.

Violations (known deviations from safe practices, procedures, standards or rules)