

To Quantify the Impact of the Existing vs. Enhanced Work Configuration of Radiation Therapy Technicians on Workload, Situation Awareness, and Performance during Pretreatment QA Tasks

Principal Investigator: Prithima Reddy Mosaly, PhD

Co-Investigators: Lukasz Mazur, PhD, Dr. Robert Adams, PhD, Lawrence Marks, MD

Division of Healthcare Engineering, Radiation Oncology Department, University of North Carolina, Chapel Hill, NC

Inclusive Dates of the Project: 7/01/2018 to 06/30/2020

Federal Project Officer: David Rodrick, PhD

Grant Number: 1R03HS025258-01A1

Acknowledgment: This study was supported by a grant from the Agency for Healthcare Research and Quality. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality.

Abstract

Purpose: To assess the impact of modifications to workflows, workspace design, and environment factors on Radiation Therapy Technicians (RTTs), workload (WL), situational awareness (SA), and performance during pretreatment quality assurance (QA) tasks in a simulated environment.

Scope: Our focus was on radiation therapy pretreatment QA tasks performed by RTTs.

Methods: Physical stressors were quantified using a rapid upper limb assessment tool (RULA). Mental WL was measured at the end of each simulated assessment subjectively using the NASA Task-Load Index (NASA-TLX) and objectively using eye-tracking methods. SA was measured using situation awareness and review technique (SART). Performance was quantified objectively using a time-out component (adherence to standard QA tasks) and error detection and procedural compliance (identification and documentation of safety/quality concerns). Finally, semi-structured interviews were conducted at the end of each simulated session to gather feedback on optimal workspace configurations. Changes in WL, SA, and performance from participants who received (vs. did not receive) enhanced workplace configurations were compared using appropriate statistical analysis.

Results: The current workplace configuration supported by 2-RTT workflow indicated increased mental WL but improved SA and performance on the time-out component ($p < 0.01$). However, the 3-RTT workflow indicated poor performance and higher RULA scores despite lower mental WL measures. The enhanced configuration reduced physical stressors (RULA; $p < 0.01$) and resulted in a higher rate of time-out compliance ($p = 0.01$). Overall, the current configuration with a 3-RTT workflow has a lower performance score and increased physical stressors.

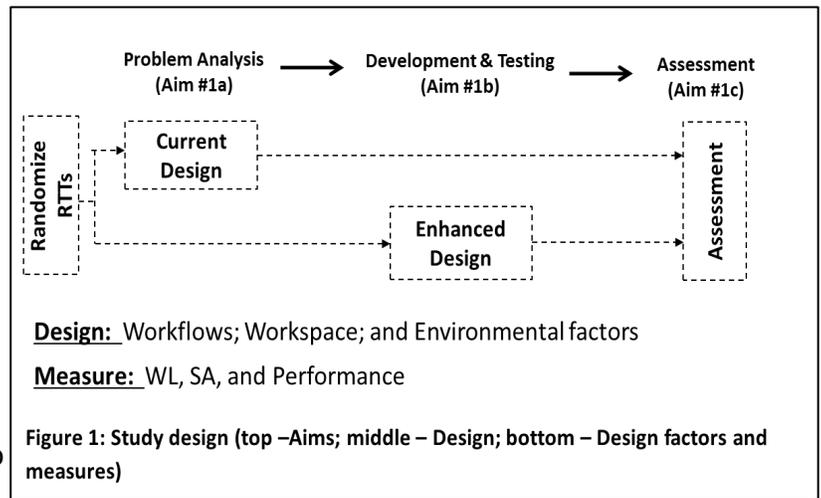
Key words: workplace configuration, mental workload, physical workload, situation awareness, performance.

a) Purpose

To assess the impact of modifications to workflows, workspace design, and environment factors on RTTs' WL, SA, and performance during pretreatment QA tasks in a simulated environment (see Figure 1).

2. Scope The present proposal addressed tasks within radiation oncology. Our focus was on radiation therapy pretreatment QA tasks performed by the RTTs. Invitations to participate in the research study were sent

to all RTTs in the UNC healthcare system and Duke University Health System while clearly stating the need for experience with treatment delivery systems as related to our simulated scenarios. All participants were incentivized to participate with a \$100 gift card. Final selections were made based on participants' ability to conduct simulated scenarios and availability to participate in the study during designated weeks for data collection.



All studies were conducted in our laboratory, a 300-square-foot dedicated room for emulating the RTTs' real clinical workspace and treatment delivery system to perform simulated pretreatment QA and treatment delivery tasks. The laboratory is divided into 2 sections; researcher station and participant station. The research and subject stations are separated via a 1-way see-through glass, and the communication to the subject is made using a 2-way microphone. The participant station is equipped with a workstation that closely emulates the real clinical environment. The participant's workstation is adjustable and includes a configurable computer monitor, a keyboard, and a computer mouse (exactly what the subjects use in the real clinic), thus increasing the fidelity of our study. The researcher's workstation allows recording and analyzing of the data from the experiments in real time. The lab is equipped with an Elekta Emulator© that allowed for the simulation of the RTTs' workflows at the treatment console, adjustable desks, and chairs for manipulating the ergonomic design of the workspace as well as a Sensomotor Inc© (SMI) goggle-based eye tracker that allowed for the collection of physiological measures of eye movements, including pupillary dilation and blinks. A Sony 64-GB HD video camera (Sony Electronics Inc©) was used to collect video recordings of RTTs' postures during scenarios.

3. Specific aim #1a: To quantitatively assess a 2-RTT vs. 3-RTT workflow on WL, SA, and performance during pretreatment QA tasks in current-standard configurations in the simulated environment.

3.1 Methods

3.1.1 Study participants: Seven RTTs [females=5, males=2; experience: >3 years] from 1 academic institution participated in this part of the study.

3.1.2 Procedure

Each participant was instructed to complete routine radiation therapy delivery tasks, including the time-out procedure, on 4 of 8 randomly assigned scenarios while working with a colleague RTT (an actor) to emulate real clinical scenarios. The time-out procedure was carried out at the console before initiating treatment. Table 1 describes the 8 scenarios and embedded errors for each scenario.

Table 1: Description of simulated scenarios and corresponding embedded error in each scenario

Scenario	Description	Embedded Error
1.	A patient with a pacemaker requiring radiation to <u>pelvis</u> who needs to have MOSFET placed for daily monitoring of dose to the pacemaker	Notice the pacemaker note and request for physics to place MOSFET
2.	A patient with a pacemaker requiring radiation to <u>left inguinal</u> who needs to have MOSFET placed for daily monitoring of dose to the pacemaker	Notice the pacemaker note and request for physics to place MOSFET
3.	A patient requiring treatment to the <u>left breast</u> involving multileaf collimator (MLC)	Incorrect/open MLC shape
4.	A patient requiring <u>pancreatic treatment</u> involving multileaf collimator (MLC)	Incorrect/open MLC shape
5.	A patient requiring <u>prostate treatment</u> with variable monitoring units (MUs)	One abnormally high MU
6.	A patient requiring <u>left breast treatment</u> with variable monitoring units (MUs)	One abnormally high MU
7.	A patient requiring <u>left breast treatment</u> and has incorrectly labeled treatment site/field name	Incorrectly labeled treatment side/field
8.	A patient requiring <u>esophagus treatment</u> and has incorrectly labeled treatment site/field name	Incorrectly labeled treatment side/field

A multidisciplinary team of medical dosimetrists, RTTs, and human factor engineers designed these scenarios based on errors reported in the literature and past incidents submitted to our department's incident reporting system. Embedded errors were carefully constructed to be rare, but realistic. No time limit was imposed on the participants to complete the scenarios.

Simulated assessments were performed using an emulator and workstations that closely replicated RTTs' typical working environment and consisted of 2 vs. 3 collocated information display configurations (Figure 2). For the 2 displays, these were the (1) electronic medical record and (2) treatment delivery information (assuming that the 3rd RTT would monitor the 3rd [not collocated] display with a video feed of a patient undergoing treatment). For the 3 displays, these were the (1) electronic medical record, (2) treatment delivery information, and (3) a video of a patient undergoing treatment. In this configuration, 1 RTT (study participant) would assume responsibility of monitoring all 3 displays.

One actor (experienced RTT) acted as the second RTT with the participants in all scenarios. Two current-configuration workflows were tested: 2 RTTs vs. 3 RTTs. In the 2-RTT (participant and actor) workflow, the participants would assume responsibility of monitoring all 3 displays [(1) electronic medical record, (2) treatment delivery information, and (3) video of patient undergoing treatment], whereas in the 3-RTT (participant, actor, virtual RTT) workflow, the participant would assume the responsibility of monitoring only 2 displays [(1) electronic medical record and (2) treatment delivery information], assuming that the 3rd RTT would monitor the 3rd (not collocated) display with a video feed of a patient undergoing treatment.

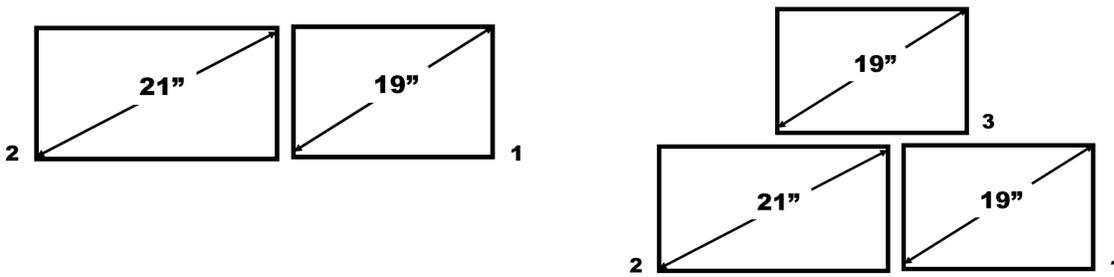


Figure 2: Left: 2-display configuration. Right: 3-display configuration

3.1.3 Data collection: In the current configuration (Figure 1; left-side pictures) emulating the real clinical setting, the desk was set to a fixed standard height of 30 inches, with the keyboard placed at 1 of 2 heights based on participants' preferences. The keyboard was placed under the desk or on the table surface, and the chair height was set to 18 inches. The 2 monitors were mounted on the table surface at a height of 47 inches, and the treatment console on the table surface was set to a height of 30 inches. Interruptions ranging from 3-5 were randomly assigned during the scenarios emulating a real clinical environment. All scenarios consisted of embedded errors (Table 1). Each participant performed 4 simulated sessions: 2 sessions each for the 2-RTT vs. 3-RTT workflows.

3.1.3.1 Physical stressors

Studies indicate that physical stressors can be improved by enhancing work-related postures.¹⁻³ McAtamney and Corlett⁴ developed an objective method (Rapid Upper Limb Assessment – RULA) for the investigation of physical stressors. RULA provides a quick assessment of postural stressors by evaluating postures on 6 anatomical regions (neck, trunk, legs, upper arms, lower arms, and wrists), force, and repetition. First, the posture of the upper limbs (upper arms, lower arms, wrist posture, and wrist twist) is determined to obtain score A. Then, postures of the neck, trunk, and legs are determined to obtain score B. Next, muscle use and force ratings are added to scores A and B to obtain scores C and D, respectively. Finally, a grand score is computed using the guidelines described by McAtamney and Corlett⁴ via a score comparison matrix.

Application: In the current study, the recorded videos were analyzed to extract and rate RTTs' postures for each simulated scenario separately. An experienced human factors researcher performed the analysis and generated a RULA score for each scenario.

3.1.3.2 Quantification of workload

Subjective WL: The National Aeronautical and Space Administration's Task Load Index (NASA-TLX)⁵ is a widely used tool for measuring subjective mental WL and has been used and validated in various domains, including radiation oncology.⁶⁻⁷ The tool measures perceived WL on 6 dimensions of WL (mental demand, physical demand, temporal demand, frustration, effort, and performance). It requires participants to perform 15 pairwise comparisons between dimensions to derive their relative weights followed by scoring each dimension on a 0- to 100-point rating scale (0=low, 100=high) based on their performance of the task under analysis. The ratings are then combined to calculate a measure of participants' mental WL as a composite NASA-TLX score ranging from 0 (low mental WL) to 100 (high mental WL). NASA-TLX scores ≥ 50 -55 have been associated with reduced performance in numerous settings,⁸⁻¹¹ including radiation oncology.^{6,7,12-13}

Application: In the current study, NASA-TLX was administered to measure mental WL at the end of each simulated scenario. RTTs were asked to rate the 6 dimensions on a paper, with verbal anchoring of each dimension, followed by the 15 pairwise comparisons. The composite NASA-TLX score for each simulated scenario was calculated based on the recommendation described above.

Objective WL: Changes in pupil diameter and blink frequency have been suggested to provide an objective (physiological) measure of mental WL.¹⁴⁻¹⁵ An increase in pupil diameter or task-evoked pupillary response (TEPR) is a subtle change in pupil size that indicates cognitive WL. Small changes

in TEPR ($\leq 0.05\text{mm}$) are involuntary and are associated with a broad range of cognitive processing that is characterized as cognitive load.¹⁶ An increase in TEPR and a decrease in blink frequency have been associated with the state of increased mental WL and a suboptimal level of performance in many basic experiments and applied studies.¹⁶⁻²⁰

Application: In the current study, data on pupil diameter and blinks were collected throughout the simulated scenarios using an SMI goggle-based eye tracker at 60Hz. The RTTs' baseline pupil data were collected by asking them to fixate on a red button located on the treatment console for 10 seconds before the start of the simulated scenario. The average baseline pupil data was then subtracted from the task pupil data (collected throughout the simulated scenarios) to obtain the task-evoked pupillary response.¹⁶ The raw eye-tracking data coded blinks or eye closures as '0'. For a continuous count of 10 to 25, '0' indicates a blink, because the average duration of a human blink may range between 100 and 400 milliseconds (for a 60-Hz frequency).¹³ A visual inspection was performed to confirm a blink vs. loss of data and identify potential outliers for validity. Potential outliers in the raw data were discarded and linearly interpolated prior to the data analysis.¹⁶⁻¹⁸ The number of blinks per minute was considered as an individual's blink rate.

3.1.3.3 Situation Awareness

The situation awareness rating technique (SART)²¹ is a subjective rating technique that measures participants' SA. SART uses 10 dimensions categorized under 3 broad categories -- demand on attentional resources (D), availability of attentional supply (S), and understanding of the situation (U) -- to measure operator SA. 'D' is measured using 3 dimensions: instability of situation (labeled as 1), variability of situation (labeled as 2), and complexity of situation (labeled as 3). 'S' is measured using 4 dimensions: arousal (labeled as 4), spare mental capacity (labeled as 5), concentration (labeled as 6), and division of attention (labeled as 7). 'U' is measured using 3 dimensions: information quality (labeled as 8), information quantity (labeled as 9), and familiarity (labeled as 10). Each dimension is rated by the participant post trial on a 7-point rating scale (1=low, 7=high) for each task under analysis. The ratings are summed to obtain the scores for each category (i.e., $U = \text{sum of ratings of dimension 1, 2, and 3}$; $S = \text{sum of ratings of 4, 5, 6, and 7}$; $D = \text{sum of ratings 8, 9, and 10}$). The global SART score is calculated as a score of the participant SA using the following formula: $SA = U - (D + S)$. *Application:* At the end of each scenario, the RTTs were asked to rate the 10 dimensions of SART on a paper, with verbal anchoring of each dimension. These ratings were combined to first calculate the scores of the 3 broad categories (U, D, S). Then, the global SART score was computed using the formula described above.

3.1.3.4 Performance

Performance was measured using 2 measures: (1) time-out compliance and (2) error detection and procedural compliance. The time-out compliance was calculated by counting number of relevant time-out components 'not missed' (conducted properly) and dividing by total relevant time-out components for each scenario (e.g., patient name, treatment site, fraction #, imaging, accessories/devices, shifts, special instructions, pacemaker-MOSFETs placement). The error detection and procedural compliance was calculated as the average score of error detection (0 denotes no detection, 1 denotes correct detection), communication (0 denotes no communication, 1 denotes communication of information that is incorrect or missing), and documentation of errors (0 denotes no documentation of error and/or missing information, 1 denotes proper documentation of error and/or missing information).

3.1.3.5 Semi-structured Interviews

Pre- and post-experiment semi-structured interviews were conducted at the beginning and the end of the experiment session for each RTT. The interview consisted of a body diagram asking RTTs to circle the anatomical regions where they felt pain or stressors. The interview also consisted of a few open-ended questions asking RTTs to comment about (1) the workspace configurations with respect to their experience of physical stressors as indicated on the body diagram; (2) any experienced suboptimal WL, SA, frustration, and temporal demand during the study; (3) their perception of key aspects of an

'optimal' workspace configuration for RTTs; and (4) any closing comments that they want to bring to researchers' attention.

Increased physical fatigue levels from clinical tasks may impact the study performance (i.e., carryover effect). Therefore, all RTTs were asked to rate their perceived fatigue level using the Borg CR-10 scale²² at the beginning of the study. Borg CR-10 is a 10-point continuous scale ranging from 0 to 10, with the first verbal anchor at 0 (no fatigue at all) going to 10 (maximal fatigue).

3.1.4 Data analysis

All measures were considered as continuous variables. The data were tested for normality (Kolmogorov-Smirnova) and an equal variance test (Levene's Test of Equality of Error Variances) for building analysis of variance (ANOVA). One-way ANOVA was used to test the effect of workflows on measures of physical stressors, mental WL, SA, and performance. An alpha of 0.05 was used for an acceptable type-I error for statistical significance. All analyses were conducted in JMP 13 software (with the alpha level set to 0.05 for significance testing).

3.2. Results: The summary of descriptive statistics is presented in Table 2.

Table 2: Descriptive Statistics of Current Configuration for 2 RTTs vs. 3 RTTs

Measures		Current	
		2 RTTs	3 RTTs
Physical Stressor	RULA	3.2 (0.6)	5 (1.3)
Subjective WL	NASA-TLX	37 (18)	22 (10)
	<i>Mental Demand</i>	56 (32)	33 (20)
	<i>Physical Demand</i>	22 (17)	10 (7)
	<i>Temporal Demand</i>	26 (24)	18 (13)
	<i>Performance</i>	27 (27)	18 (12)
	<i>Effort</i>	46 (32)	26 (18)
	<i>Frustration</i>	35 (34)	11 (5)
Objective WL	TEPR (mm)	0.16 (0.10)	0.13 (0.08)
	Blink Rate (blinks/min)	9 (4)	11 (2.4)
SA	SART	19 (3)	17 (3)
	<i>Demand</i>	2.6 (0.9)	3.3 (1.4)
	<i>Supply</i>	3.6 (0.7)	3.9 (1.0)
	<i>Understanding</i>	6.1 (0.5)	5.3 (0.5)
Performance	Time-out Component (%)	82 (18)	50 (14)
	Error Detection and Compliance (%)	48 (36)	15 (28)

3.2.1 Physical Stressors

There was a significant effect of the 2-RTT vs. 3-RTT workflow on RULA scores, indicating that the 3-RTT workflow has the most suboptimal postures compared with the 2-RTT workflow ($F_{(1,26)}=26$, $p<0.0001$).

3.2.1 Workload

A significant effect of the 2-RTT vs. 3-RTT workflow on NASA-TLX ($F_{(1,26)}=7.5$, $p=0.01$) and 3 of the 6 dimensions [mental demand ($F_{(1,26)}=4.8$, $p=0.03$), physical demand ($F_{(1,26)}=6.0$, $p=0.02$), and frustration ($F_{(1,26)}=6.6$, $p=0.02$)] were observed. The 2-RTT workflow had a significantly higher workload compared with the 3-RTT workflow. No significant effect was observed on objective measures of WL ($p>0.05$).

3.2.2 Situation awareness

No significant effect of the 2-RTT vs. 3-RTT workflow on SART scores and the 3 categories was observed ($p>0.05$).

3.2.3 Performance

Both performance measures, time-out component ($F_{(1,26)}=26$, $p<0.0001$) and error detection and compliance ($F_{(1,26)}=6.9$, $p=0.01$), were significantly affected by the 2-RTT vs. 3-RTT workflow, indicating that the 2-RTT workflow has the most optimal performance compared with the 3-RTT workflow.

3.2.4 Semi-structured Interviews

The average Borg CR-10 rating 2.8 (1.6) and all RTTs (7/7) indicated that they experience physical stressors in 1 or more anatomical locations (primarily on the neck, lower back, shoulder blades) indicated on the body diagrams. The RTTs commented that their workspaces need ergonomic improvements. One RTT said “Go look at any machines. No legroom due to monitors (CPUs) and wires. Prefer less phone interruptions.” These findings support ergonomic enhancements to workspace settings of RTTs.

3.2.5 Discussion

The 2-RTT workflow increased the subjective mental WL; however, the performance was optimal, and RTTs were able to maintain near-neutral postures compared with the 3-RTT workflow. These findings support the finding that the 2-RTT workflow is optimal; however, measures should be taken to minimize mental WL. Further research is needed to explore ways to optimize WL (e.g., simulation-based training, electronic checklists, etc.) while maintaining higher levels of performance. Although the physical stressors are lower compared with the 3-RTT workflow, the 2-RTT workflow in the current configuration has an average RULA score range of 3.2 (0.6) to 5 (1.3), indicating that there is a mild to moderate possibility of developing work-related musculoskeletal disorders in the future. The semistructured interviews results indicated that all RTTs experienced stressors in 1 or more body part (indicated on the body diagram) in their work and indicated that the workspaces need improvement. Hence, we improved the workspaces based on the RTT comments and the best practices and standards established in the literature. In the next phase, we designed and tested the enhanced workspaces on WL, SA, and performance.

4. Specific aim #1b: To develop and quantitatively assess RTTs' WL, SA, and performance during pretreatment QA tasks in an enhanced configuration in the simulated environment.

4.1 Methods

4.1.1 Participants and setting

Seven RTTs [females=4, males=3; experience: >3 years] from 2 academic institutions participated in this part of the study. In the enhanced configuration, we ergonomically optimized the desk and chair heights

as well as location and viewing angles of monitors based on the feedback we received from RTTs in Aim 1a and the ergonomics standards established by Sanders and McCormick.²³ The enhancements were also calibrated based on participants' physical characteristics (RTTs' height range: 5 feet 4 inches to 6 feet; adjusted chair height range: 16-22 inches; and desk height range: 25-30 inches; standing desk height range: 40-44.5 inches). The keyboard and treatment console were set on the table. The 2 monitors were mounted on the monitor arm for height (46-67 inches) and tilt (10-20 degrees) adjustments (Figure 1).

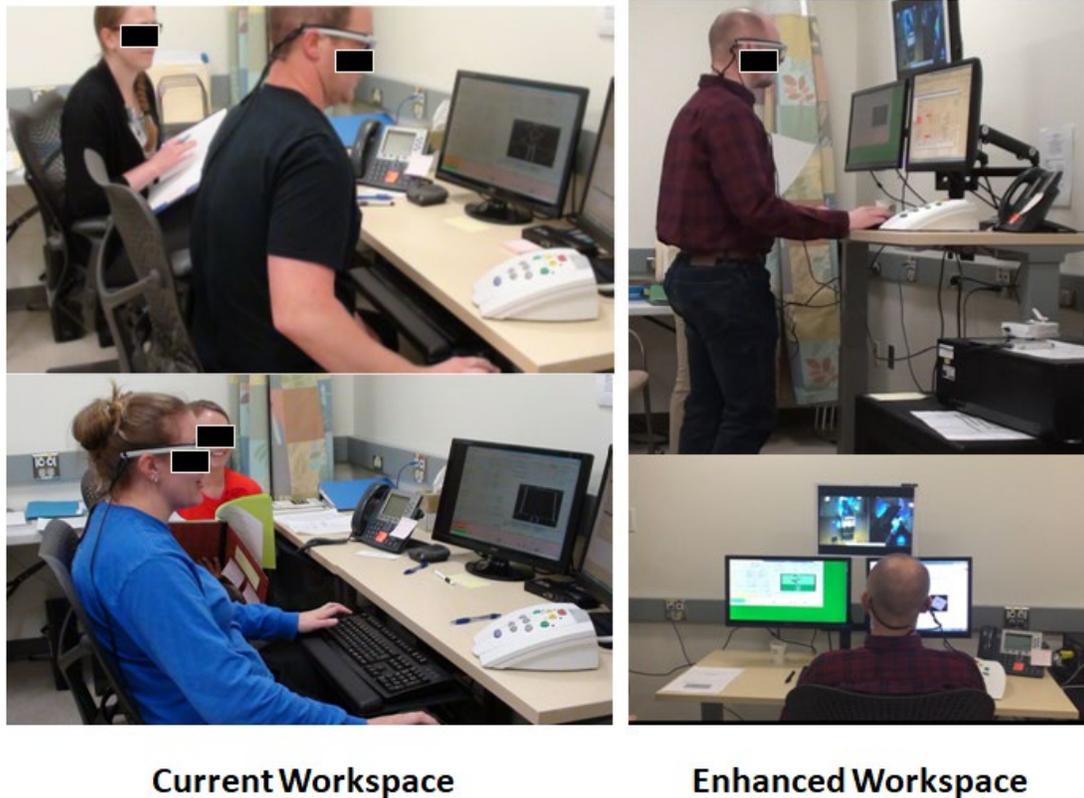


Figure 3: Workspace configuration setting. Left: current setting; Right: enhanced setting (top: elevated seating/standing; bottom: regular seating)

4.1.2 Procedure: Same as section 3.1.2

4.1.3 Data collection: Same as section 3.1.3

4.1.3.1 Physical stressors: Same as section 3.1.3.1

4.1.3.2 Quantification of workload: Same as section 3.1.3.2

4.1.3.3 Situation awareness: Same as section 3.1.3.3

4.1.3.4 Performance: Same as section 3.1.3.4

4.1.3.5 Semi-structured interviews: Same as section 3.1.3.5

4.1.4 Data analysis: Same as section 3.1.4

4.2. Results: The summary of descriptive statistics is presented in Table 3.

Table 3: Descriptive Statistics of Enhanced Configuration for 2 RTTs vs. 3 RTTs

Measures		Enhanced	
		2 RTTs	3 RTTs
Physical Stressor	RULA	3.3 (0.5)	3.4 (0.5)
Subjective WL	NASA-TLX	35 (20)	35 (19)
	<i>Mental Demand</i>	21 (29)	48 (26)
	<i>Physical Demand</i>	23 (18)	19 (18)
	<i>Temporal Demand</i>	24 (19)	23 (20)
	<i>Performance</i>	37 (32)	33 (30)
	<i>Effort</i>	39 (26)	25 (24)
	<i>Frustration</i>	37 (33)	39 (31)
Objective WL	TEPR (mm)	0.2 (0.15)	0.16 (0.14)
	Blink Rate (blinks/min)	10 (4)	9 (4.8)
SA	SART	19 (7)	19 (6)
	<i>Demand</i>	2.8 (1.2)	3.1 (1.3)
	<i>Supply</i>	4.2 (0.9)	4.1 (0.8)
	<i>Understanding</i>	6.4 (0.9)	6 (1.4)
Performance	Time-out Component	79 (16)	80 (15)
	Error Detection and Compliance	69 (40)	33 (47)

4.2.1 Physical Stressors

No significant effect in the 2-RTT vs. 3-RTT workflow on RULA scores ($p > 0.05$) was observed.

4.2.1 Workload

No significant effect of the 2-RTT vs. 3-RTT workflow on subjective and objective WL measures ($p > 0.05$) was observed.

4.2.2 Situation Awareness

No significant difference in SART scores between the 2-RTT vs. 3-RTT workflow ($p > 0.05$) was noted.

4.2.3 Performance

A significant effect of the 2-RTT vs. 3-RTT workflow on error detection and compliance was observed ($F_{(1,26)}=4.2$, $p=0.04$), indicating that the 2-RTT workflow has improved performance compared with the 3-RTT workflow.

4.2.4 Semi-structured Interviews

The average Borg CR-10 rating 2.4 (1.5) and most RTTs (6/7) indicated that they experience physical stressors in 1 or more anatomical locations (primarily on the lower back, shoulders, and feet) indicated on the body diagrams. The RTTs commented that their workspaces need ergonomic

improvements. The RTTs indicated that the lab setting was more comfortable than their real workstations, because their real workstations are not flexible or adjustable. They also commented that a comfortable chair, with flexibility to change to sit or stand during work, and the correct level of monitor heights are important to their work.

4.3 Discussion

Overall, the study findings indicated that the RTTs experienced similar levels of physical stressors, and optimal levels of WL and SA, in the enhanced configuration in both 2-RTT and 3-RTT workflows. However, the performance was optimal in the 2-RTT workflow compared with the 3-RTT workflow. These findings suggest that further testing and validation are required to confirm the current findings.

5. Specific Aim #1c: To quantitatively compare the impact of current-standard vs. enhanced configurations on RTTs' WL, SA, and performance during pretreatment QA tasks.

5.1.1 Participants and setting: The data from Aims 1a (n=7) and 1b (n=7) were combined to perform the data analysis. Same as section 3.1.1

5.1.2 Procedure: Same as section 3.1.2

5.1.3 Data collection: Same as section 3.1.3

5.1.3.1 Physical stressors: Same as section 3.1.3.1

5.1.3.2 Quantification of workload: Same as section 3.1.3.2

5.1.3.3 Situation awareness: Same as section 3.1.3.3

5.1.3.4 Quantification of performance: Same as section 3.1.3.4

5.1.3.5 Semi-structured interviews: Same as section 3.1.3.5

5.1.4 Data analysis: Same as section 3.1.4

5.2 Results

5.2.1 Physical stressors

There was a significant main and interaction effect of workspace configuration and 2-RTT vs. 3-RTT workflow on RULA scores ($F_{(3,55)}=18.4$, $p<0.0001$, Figure 4). The results indicated that current configuration with the 3-RTT workflow has the most suboptimal postures compared with others.

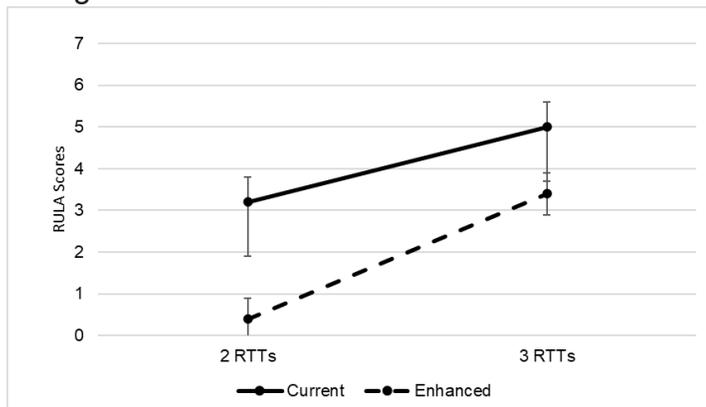


Figure 4: Interaction effect of workspace (current vs. enhanced) and workflow (2 RTTs vs. 3 RTTs) on performance measure vs. time-out component

5.2.2 Mental workload

No significant effect of workspace configuration and workflow on subjective and objective measures of WL ($p>0.05$) was observed.

5.2.3 Situation awareness

There was a significant main effect of configuration and workflow on understanding (U) of the situation ($F_{(3,55)}=3.8$, $p=0.01$), indicating that the 2-RTT workflow and enhanced configuration had improved situation understanding of the situation compared with others. No significant effect of workplace configuration and configuration on global SART scores and demand (D) and supply (S) categories of the SART were observed ($p>0.05$).

5.2.4 Performance

There was a significant main and interaction effect of workspace configuration and workflow on the time-out component ($F_{(3,55)}=13$, $p<0.0001$), and there was a significant main effect of workflow on error compliance ($F_{(3,55)}=4.8$, $p=0.005$, Figure 5). The results indicated that the 3-RTT workflow in current configuration setting has the most suboptimal performance on the time-out component compared with others; the 3-RTT workflow had the most suboptimal error detection and compliance compared with the 2-RTT workflow.

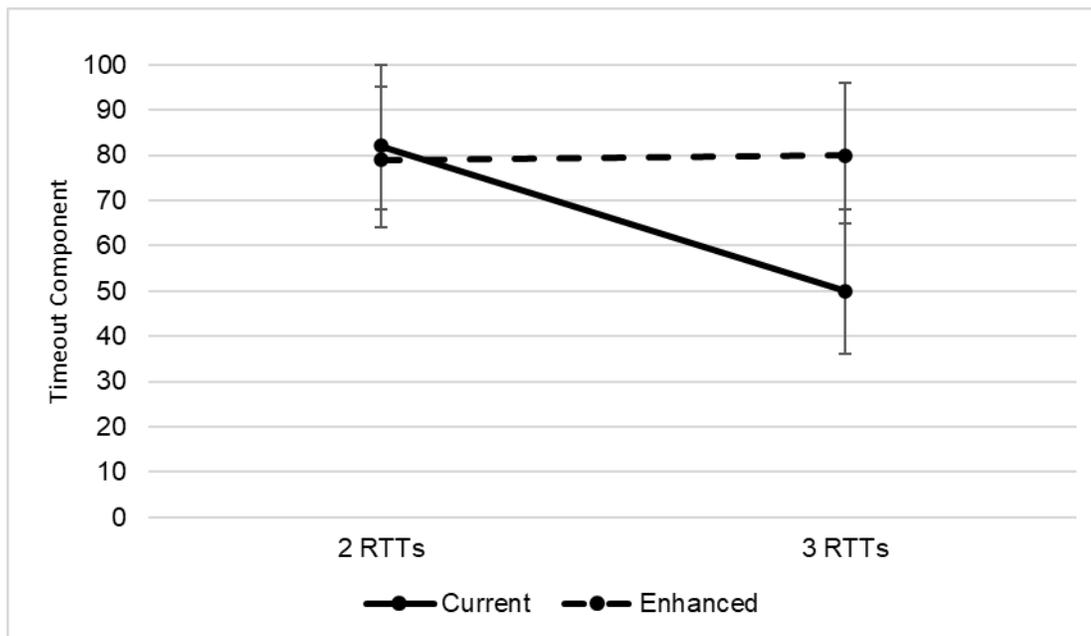


Figure 5: Interaction effect of workspace (current vs. enhanced) and workflow (2 RTTs vs. 3 RTTs) on performance measure vs. time-out component

5.2.5 Semi-structured Interviews

There was no significant statistical difference between RTTs randomized to the current configuration ($M=2.8$, $SD=1.5$) vs. those randomized to enhanced configuration ($M=2.4$, $SD=1.5$) on the physical fatigues measure via the Borg CR-10 rating. RTT feedback is summarized in Table 4.

Table 4: Summary of the open ended questionnaire administered to participants to get their feedback on current work configuration and assess their requirement for ideal workspace configuration

Current Workstation	Ideal Workstation
Need a taller desk, more back support, possibly more cushioning in chair	Leg room: "Go look at any machines. No legroom due to monitors (CPUs) and wires." Less phone interruptions
Mouse location low, would make wrist hurt long term	Enough space; the ability to stand and/or sit
"I don't like the keyboard drawer or the mouse area (too small and hits my legs) . Not enough computer space around Elekta box - monitor is too small."	Level of screen, lightning, comfort - sitting or standing
Keyboard was lower than I'd like, mouse moved higher for my comfort	Having computers in correct order of treatment process and having room for all equipment that is easily accessible especially for emergent situations
Better cushioning for floors would make it easier on the feet and knees	Eyes on computer and patient at the same time
Mouse location is poor, Elekta console location is also poor	Rotational positioning/standing & sitting to change position with a comfortable mat for standing and as few monitors as possible to prevent eye fatigue
The in-room monitor versus the treatment monitors are not in the same field of view	Adjustable desk height with adjustable monitors and chairs
	A sit-and-stand station with the treatment delivery monitor and the in-room patient monitor in the same field of view; 2 monitors for treatment only; a 3 rd monitor away from the therapist driving

5.3 Discussion

The current aim assessed the effects of workspace configuration and workflows on RTTs' physical stressors, mental WL, SA, and performance during routine treatment delivery tasks in a simulated environment. First, the study results suggest that the 2-RTT workflow in enhanced configurations has lower physical stressors and maintains an optimal level of SA and performance compared with the 3-RTT workflow. The 2-RTT workflow resulted in the most optimal performance, and we presume that the RTTs have learned to maintain higher levels of SA and performance during pretreatment QA tasks, because the 2-RTT workflow represent a real clinical scenario.

Second, an ergonomically optimized workspace configuration significantly reduces physical stressors on RTTs as quantified by RULA. The current workspace configuration resulted in suboptimal postures of RTTs' neck, lower back, shoulders, and wrists. The findings from this study suggest that current workspace designs could present a potential risk for developing work-related injuries in RTTs over time and, thus, should be improved. During the post-experiment interviews in the current study, 70% of RTs (10 of 14) indicated that they experience physical pain in at least 2 locations of their body at the end of each working day (e.g., especially in the neck and lower back). All the participants (14 of 14) indicated that ergonomically optimized workspaces would better fit their individual needs.

Third, the study results indicated no significant differences in the subjective mental WL, as quantified by the global and individual dimensions of the NASA-TLX scores, and objective mental WL, as quantified by the eye metrics (pupil diameter and blink rate) between the 2 configurations. It is certainly possible that the magnitude of the configuration changes made in the current study was not significant enough to generate a meaningful change in RTTs' perceptions of mental WL. This was later confirmed during

our post-experiment interviews with RTTs. That is, the RTTs considered all scenarios, regardless of the configuration changes, equally challenging and similar to what they experience in the actual clinical setting.

Fourth, the study results indicated no significant differences in SA as quantified by the global SART score. Nevertheless, there were significant improvements in the level of understanding of the situation (U) in both enhanced configuration and the 2-RTT workflow setting. From the post-experiment interviews, we learned that this result was mostly attributed to a more optimal location and viewing angles of monitors based on RTTs' physical characteristics. However, to our knowledge, there is no prior study that assessed the effects of physical stressors on SA in the radiation therapy domain; hence, further evaluation is needed to confirm our results.

Finally, the study results indicated a significant effect of enhanced workspace configuration and the 2-RTT workflow on improved scores of time-out compliance and the 2-RTT workflow on the improved error detection and procedural compliance scores. From the post-experiment interviews, we learned that the enhanced configurations have improved overall levels of information perception and thus spearheaded better performance on specific time-out components. Furthermore, the 2-RTT workflow represents the real clinical workflow; hence, the (primary) RTTs are trained to take the complete responsibility of overall safety and quality of the pretreatment QA tasks.

6. Limitations

6.1 Sample size and manipulation of environment factors (i.e., interruptions)

Due to changes in the organization (increase in patient load, installation of new treatment machines, administrative changes in the workflow, etc.), there was a significant delay in the recruitment of the study participants within the anticipated timeline. Hence, we were unable to recruit 15 participants for Aim #1a and Aim #1b. However, we were able to meet the required number of RTTs for the study (total of 14 RTTs participated in the study). Also, due to the time constraint, the above changes have also resulted in not including the environmental factors as one of the study variables during the study.

6.2 Study limitations

- a) First, study findings are based on 14 RTTs from 2 large academic medical centers. Although the sample size was relatively small, all RTTs in the study were experienced in performing the RTT pretreatment QA and delivery tasks, thus representing a large population of RTTs.
- b) Second, we used the emulator and the console in a laboratory setting with no direct connection to the real treatment machine. This may have confounded our results, specifically in terms of sensitivity in identifying the difference in mental WL, SA, and performance with changes in workspace configurations (in specific Aim #1c).
- c) Third, there was no real patient involved in the simulated scenarios. Although the emulator has the capacity to emulate a real clinical environment, the RTTs may not be exerting the same amount of cognitive resources nor maintaining the same high levels of SA as they would otherwise in a real clinical setting. It is a recognized drawback in virtually all simulated settings.
- d) Fourth, overall fatigue from clinical work and other psychosocial factors could have affected results. In the current study, RTTs reported relatively low levels of fatigue. Further, psychosocial and individual factors are also recognized threats for musculoskeletal disorders.²⁴ However, the current study aims to target workplace configuration to enhance postures for design improvements based on the best practices.²³ Further studies are needed to assess the multifactorial approach in developing work-related injuries in RTTs.
- e) Finally, techniques employed to calculate eye-based metrics are challenging. For example, pupillary dilations are sensitive to and affected by changes in illumination of display monitors, change in

visual gaze, etc. Therefore, a visual inspection was performed to identify potential outliers in the raw eye-tracking data, which were later discarded and linearly interpolated prior to data analysis based on standards set for processing eye-tracking data. ^{16-18,24}

7. Conclusions

Limited research has been done previously to quantify the impact of ergonomic enhancements to workspaces and workflows on physical stressors, mental WL, SA, and performance of RTTs. To our knowledge, this study is the first of its kind in the radiation therapy domain to assess the effects of workspace configuration on physical stressors, mental WL, SA, and performance. The current study results suggest that ergonomically designed workspaces and careful consideration of workflows may reduce physical stressors and improve RTTs' time-out compliance.

8. References:

1. Andersson, G., Hagman, J., Talianzadeh, R., Svedberg, A. and Larsen, H.C., 2002. Effect of cognitive load on postural control. *Brain research bulletin*, 58(1), pp.135-139.
2. Ehrenfried, T., Guerraz, M., Thilo, K.V., Yardley, L. and Gresty, M.A., 2003. Posture and mental task performance when viewing a moving visual field. *Cognitive Brain Research*, 17(1), pp.140-153.
3. Liao, M.H. and Drury, C.G., 2000. Posture, discomfort and performance in a VDT task. *Ergonomics*, 43(3), pp.345-359.
4. McAtamney, L. and Corlett, E.N., 1993. RULA: a survey method for the investigation of work-related upper limb disorders. *Applied Ergonomics*, 24(2), pp.91-99.
5. Hart, S.G. and Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139-183). North-Holland.
6. Mazur, L.M., Mosaly, P.R., Hoyle, L.M., Jones, E.L., Chera, B.S. and Marks, L.B., 2014. Relating physician's workload with errors during radiation therapy planning. *Practical Radiation Oncology*, 4(2), pp.71-75.
7. Mosaly, P.R., Mazur, L.M., Jones, E.L., Hoyle, L., Zagar, T., Chera, B.S. and Marks, L.B., 2013. Quantifying the impact of cross coverage on physician's workload and performance in radiation oncology. *Practical Radiation Oncology*, 3(4), pp.e179-e186.
8. Calkin, B.A., 2007. Parameters affecting mental workload and the number of simulated UCAVS that can be effectively SUPERVISED (Masters Thesis, Wright State University). Source retrieved on Feb 24 2020 from <https://pdfs.semanticscholar.org/4b5a/5869798c6078dd478a5355ce41426a1f22c5.pdf>
9. Yurko, Y.Y., Scerbo, M.W., Prabhu, A.S., Acker, C.E. and Stefanidis, D., 2010. Higher mental workload is associated with poorer laparoscopic performance as measured by the NASA-TLX tool. *Simulation in Healthcare*, 5(5), pp.267-271.
10. Mazur, L.M., Mosaly, P.R., Moore, C., Comitz, E., Yu, F., Falchook, A.D., Eblan, M.J., Hoyle, L.M., Tracton, G., Chera, B.S. and Marks, L.B., 2016. Toward a better understanding of task demands, workload, and performance during physician-computer interactions. *Journal of the American Medical Informatics Association*, 23(6), pp.1113-1120.
11. Mosaly, P.R., Mazur, L.M., Yu, F., Guo, H., Derek, M., Laidlaw, D.H., Moore, C., Marks, L.B. and Mostafa, J., 2018. Relating task demand, mental effort and task difficulty with physicians' performance during interactions with electronic health records (EHRs). *International Journal of Human-Computer Interaction*, 34(5), pp.467-475.
12. Mazur, L.M., Mosaly, P.R., Hoyle, L.M., Jones, E.L. and Marks, L.B., 2013. Subjective and objective quantification of physician's workload and performance during radiation therapy planning tasks. *Practical Radiation Oncology*, 3(4), pp.e171-e177.
13. Mazur, L.M., Mosaly, P.R., Jackson, M., Chang, S.X., Burkhardt, K.D., Adams, R.D., Jones, E.L., Hoyle, L., Xu, J., Rockwell, J. and Marks, L.B., 2012. Quantitative assessment of workload and stressors in clinical radiation oncology. *International Journal of Radiation Oncology* Biology* Physics*, 83(5), pp.e571-e576.
14. Beatty, J., 1982. Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological bulletin*, 91(2), p.276.
15. Poole, A. and Ball, L.J., 2006. Eye tracking in HCI and usability research. In *Encyclopedia of Human Computer Interaction* (pp. 211-219). IGI Global.
16. Beatty J, Lucero-Wagoner B. The Pupillary System. In: Cacioppo JT, Tassinary LG, Bernston GG, editors. *Handbook of Psychophysiology*. 2nd ed. Cambridge University Press; 2000. p. 142-162.
17. Marshall, S.P., 2005. Assessing cognitive engagement and cognitive state from eye metrics. *Foundations of augmented cognition*, 11, pp.312-320.
18. Iqbal, S.T., Zheng, X.S. and Bailey, B.P., 2004, April. Task-evoked pupillary response to mental workload in human-computer interaction. In *CHI'04 extended abstracts on Human factors in computing systems* (pp. 1477-1480).

19. De Waard, D. and te Groningen, R., 1996. The measurement of drivers' mental workload. Netherlands: Groningen University, Traffic Research Center.
20. East, J.A., Bauer Jr, K.W. and Lanning, J.W., 2002. Feature selection for predicting pilot mental workload: A feasibility study. *International Journal of Smart Engineering System Design*, 4(3), pp.183-193.
21. Taylor, R.M., 1990. Situational Awareness Rating Technique (SART): The development of a tool for aircrew systems design. AGARD. *Situational Awareness in Aerospace Operations*, 17, pp.23-53.
22. Borg, G.A., 1982. Psychophysical bases of perceived exertion. *Medicine & Science in Sports & Exercise* 14, 377-381.
23. Sanders, M.S. and McCormick, E.J., 1993. Applied anthropometry, work-space design and seating. *Human Factors in Engineering and Design*, 7.

9. List of Publications

We reported results in:

- 1. Prithima R Mosaly, Robert Adams, Gregg Tracton, John Dooley, Karthik Adapa, Joseph K Nuamah, Lawrence B Marks, Lukasz M Mazur. Impact of Workspace Design on Radiation Therapist Technicians' Physical Stressors, Mental Workload, Situation Awareness and Performance. Submitted to Practical Radiation Oncology.**
Summary: The enhanced configuration significantly reduced physical stressors and resulted in a higher rate of time-out compliance compared with current workspace configuration. Our results suggest that an ergonomically designed workspace may minimize physical stressors and improve the performance of RTTs.
- 2. Joseph K Nuamah, Prithima R Mosaly, Robert Adams, Kathik Adapa, Bisham S Chera, Lawrence B Marks, Lukasz M Mazur. Assessment of Radiation Therapy Technologist's Workload and Situation Awareness: Monitoring Two vs. Three Collocated Display Monitors. Submitted to Advances in Practical Radiation Oncology.**
Summary: The addition of the third collocated display appears to have no significant impact on RTTs' WL or SA during treatment delivery in enhanced workspace configuration. This is a noteworthy finding, suggesting that monitoring 3 displays by 1 RTT is reasonable from human factors perspective and could be considered for clinical practice.
- 3. Mosaly P, Mazur LM: Abstract #1185, submitted to AHFE 2019 International Conference, July 2019.**
Abstract: Suboptimal design of workflows, workspaces, and environmental factors likely hinders safety in Radiation Therapy (RT). Errors in RT are estimated to occur in up to $\approx 5\%$ of the $> \approx 600,000$ patients receiving RT per year in the US; with serious/lethal events occurring ≈ 1 of 1,000-10,000 patients. Flaws in workflows are recognized threats to safety in many industries, including healthcare. Similarly, physical workspace design (e.g., congested workspaces, ergonomic designs) and environmental stressors (e.g., excess noise, inefficient lighting, interruptions) are identified as factors negatively affecting human performance and patient safety. Several recent reports in the academic literature and lay press also highlight unique patient safety challenges within RT, particularly during quality assurance (QA) procedures of RT technicians (RTTs). Within the RT field, essentially all patient harm is manifest at the treatment machine, with either upstream errors being propagated or new errors being generated. In addition, the American College of Radiology for radiation oncology practice accreditation (ACR 2017) recommends 2.6 to 3.5 full-time RTTs per treatment machine, whereas, the institutes employ only 2 to 2.5 full-time RTTs due to cost burden. This further magnifies patient safety concerns, specifically during the critical task performance like QA. Thus, it is critical to do what we can to maximize the performance of RTTs, especially during pre-RT QA tasks.
Goal: Overall goal of this study was to assess the impact of modifications to workflows, workspace design, and environment factors on RTTs' physical and mental workload (WL), situation awareness (SA), and performance during pretreatment QA tasks.
Methods: Eight RTTs participated in a simulated RT pretreatment QA task in the lab environment. The workspace was set up to emulate the current (baseline) clinic configuration for workflows (2 RTTs per machine), workspace design (desk height, monitors configuration, and radiation delivery console location), and environment (1 to 2 interruptions per simulation). The physical workload was measured using Borg CR-10 scale and Rapid Upper Limb Assessment (RULA); mental workload was measured using NASA task load index (TLX); SA was measured using situation awareness and review technique (SART); and performance was measured using task-completion

time and identification of embedded errors. An open-ended questionnaire was also administered at the end of the scenario to capture participants' feedback on optimal (enhanced) configuration. We are in the process of creating an enhanced configuration and will assess (using the same procedure as described above) the impact on the WL, SA, and performance. We also will compare the results between the current and enhanced configuration.

Results: We anticipate that the enhanced configuration will minimize both physical and mental WL and improve SA and performance when compared with current configuration.

Conclusions: The study findings will not only provide guidelines to design workflows, workspaces, and environmental factors that could facilitate optimal RTT performance during pre-RT QA tasks but also help policymakers to make decisions on required staffing levels for patient safety and to expand this learning and improvements to other areas in healthcare for generalizability.

4. Mosaly P, Mazur LM: Radiation Oncology Safety Retreat at Duke University Health System on October 2019.

Abstract: To assess the impact of current vs. enhanced workspace design on RTTs' physical and mental workload (WL), situation awareness (SA), and performance during pretreatment QA tasks.

Method: 14 RTTs participated in this study. Physical stressors, MWL, SA, and performance were assessed while performing 4 simulated scenarios in an internal review board (IRB)-approved study. Physical stressors were quantified using the rapid upper limb assessment (RULA) method; MWL was measured at the end of each simulated scenario subjectively using the NASA Task-Load Index (NASA-TLX); and SA was measured at the end of the scenario using situation awareness and review technique (SART). Performance was measured objectively via assessment of compliance with pretreatment quality assurance (QA) tasks (time-out), and identification of embedded errors with appropriate follow-up actions. Pearson and Spearman's Rho (for RULA) correlation analysis was conducted between performance and physical stressors, NASA-TLX, and SA. An alpha value of 0.05 was used for statistical significance.

Results: The analysis indicated that there was a significant positive correlation between performance and NASA-TLX RULA scores ($r=-0.3$, $p=0.03$) and a significant negative correlation between performance and RULA scores ($p=-0.4$, $p<0.01$). There was a significant negative correlation between RULA and NASA-TLX ($p=-0.3$, $p=0.02$).

Discussion: The current study results indicated that a physical stressor has an impact on perceived workload and performance. Therefore, improving workspace design and workstation layout can improve RTT performance. Well-designed workspaces and workflows not only improve patient safety but also the well-being and health of care providers.

5. Mosaly P, Mazur LM: Predicting Radiation Therapy Technicians' Performance using Physical Stressors, Mental Workload and Situation Awareness. Abstract #40142, submitted to ASTRO Annual Meeting 2020.

Abstract: To assess the impact of physical stressors, mental workload (MWL), and situational awareness (SA) on Radiation Therapy Technicians' (RTTs') performance during routine treatment delivery tasks in a simulated environment.

Materials/Methods: 14 RTTs participated in this study. Physical stressors, MWL, SA, and performance were assessed while performing 4 simulated scenarios in an internal review board (IRB)-approved study. Physical stressors were quantified using the rapid upper limb assessment (RULA) method; MWL was measured at the end of each simulated scenario subjectively using the NASA Task-Load Index (NASA-TLX) and objectively throughout the scenario using eye-tracking measures for blink rate and task-evoked pupillary response; and SA was measured at the end of the scenario using situation awareness and review technique (SART). Performance was measured objectively via assessment of compliance with pretreatment quality assurance (QA) tasks (time-out) and identification of embedded errors with appropriate follow-up actions. Multiple linear regression

analysis was used to develop a model for predicting performance based on physical stressors, MWL, and SA. An alpha of 0.05 was used for statistical significance.

Results: Overall, the compliance with pretreatment quality assurance (QA) tasks (time-out) was predicted by physical stressors and MWL ($p < 0.001$; $\text{adj-R}^2 = 0.45$), indicating a strong negative correlation with both RULA ($p < 0.001$; indicating suboptimal ergonomic postures) and blink rate ($p < 0.01$; indicating suboptimal MWL).

Discussion: Suboptimal postures are known to cause work-related musculoskeletal injuries, and suboptimal levels of mental workload are known to decrease performance in many safety critical industries, including healthcare. Therefore, designing workstations that can improve RTTs' posture and minimize physical stressors and enhancing visual displays for optimal MWL and information presentation could improve RTTs' performance during QA pretreatment tasks.

Conclusions: Limited research has been done in the area of Radiation Therapy in assessing the impact of multiple factors on RTTs performance. Future research should consider various measures of physical stressors, MWL, and SA and develop models that can guide researchers in designing optimal workspace to improve RTTs performance.

10. List of Products

Based on our findings, we proposed following recommendations:

- **Recommendations #1:** A 2-RTT workflow is the most optimal workflow for pretreatment QA tasks. However, consideration should be made to optimize WL by including other interventions (e.g., simulation-based training, electronic checklists, etc.).
- **Recommendation #2:** Enhancing physical workspaces and optimizing configuration will have greater benefit in minimizing physical stressors on RTTs and may enhance SA and performance, thus improving patient safety and quality in the radiation therapy domain.