Final Report: Automated Image Analysis for the Prevention of Radiotherapy Delivery Errors

Resubmission Date 1-10-2024

Title of Project: Automated Image Analysis for the Prevention of Radiotherapy Delivery Errors

Principal Investigator: James M. Lamb, Ph.D.

Co-Investigators: Naomi Bardach M.D., Matthew Brown Ph.D., Fang-I Chu Ph.D, Daniel Low Ph.D., John P. Neylon Ph.D., Ann Raldow M.D., Tim Ritter Ph.D.

Graduate Student Researchers: Rachel Petragallo, Dishane Luximon, John Charters

Organization: University of California, Los Angeles

Inclusive Dates of Project: 9/30/2018-7/31/2023

Federal Project Officer: Darryl Gray MD

Acknowledgment of Agency Support: This work supported by Agency for Healthcare Research and Quality

Grant Award Number: 1R01HS026486

Structured Abstract

Purpose: To develop algorithms for the detection of patient positioning and patient identification in image guided radiotherapy, to apply those algorithms retrospectively to clinical image databases to detect un-reported errors, and to implement the algorithms as error-prevention system and study their impacts in a clinical environment.
Scope: The scope of this project is image-guided radiation therapy, in which patients are positioned for treatment using x-ray and cone-beam computed tomography images acquired using on-board imagers. Image guidance is used for over 90% of treatments in a modern radiotherapy clinic.
Methods: Error detection algorithms were developed using deep convolutional neural nets for all the major radiotherapy image guidance technologies: planar kilovoltage (kV) x-ray, planar megavoltage (MV) x-ray, and cone-beam computed tomography (CBCT). Algorithms were applied retrospectively to image databases of radiotherapy departments at University of California, Los Angeles, and Virginia Commonwealth University.
Results: Algorithm accuracy, characterized by Receiver Operating Characteristic (ROC) Area Under the Curve (AUC), was 0.985-0.997 for planar kV x-ray, 0.92 for MV x-ray, and 0.995 for CBCT. Application to retrospective databases yielded error rates of 0.044% [0.0022%, 0.21%] for planar kV imaging, and 0.04%+-0.02% for CBCT imaging. Of a total of 5 errors identified, 2 were previously unreported. The CBCT algorithm was implemented in a prototype clinical system that produced a summary report of alignment quality and errors that was viewable on a web interface. Structured interviews of clinical physicists were carried out to elicit barriers and facilitators to clinical use of the prototype tool.

Key Words: radiotherapy, IGRT, medical event

Purpose.

The objectives of the study are 3-fold:

1) Develop algorithms to detect mismatches between the planned and actual patient position in image guided radiotherapy, which would allow detection and prevention of delivery of radiation to the wrong part of the body, or treatment a patient with a different patient's radiotherapy plan.

2) Use these algorithms to search through clinical image databases to identify previously un-reported errors. Most if not all knowledge of error rates in radiotherapy is from error reporting. Error reporting is known to underestimate error prevalence because not all errors are detected, and barriers exist that limit reporting of known errors. Thus, this aim addresses a knowledge gap.

3) Implement the algorithms in real-time or near-real time clinical systems and study the effect of the algorithms in a clinical environment.

Scope

Approximately 1 million patients per year receive radiation therapy in the United States as part of their cancer care. Rapidly advancing technology has over the past two decades enabled treatments in which irradiated volumes are made to conform ever more tightly to tumors. Increasing conformality has resulted in better cure rates and lower side effects, but requires highly precise patient positioning with little margin for error. Patient positioning is performed by radiation therapy technologists, usually by visually aligning on-board x-ray setup images to the CT scans used to plan patients' treatments. Academic studies, public records, and works of investigative journalism have demonstrated that, despite quantitative positioning processes and extensive and rigorous quality control, human error leads to so-called never events: treatments with serious alignment errors or with the wrong patient's treatment plan. Based on never events reported at UCLA, at least 1,400 such events occur nationally per year.

Though rare, radiotherapy never events have potentially devastating consequences, and reducing their occurrence is strongly motivated. Radiotherapy is already subject to intensive quality control procedures, so further reduction is likely best achieved through automation in order to avoid additional burden on an already labor-intensive workflow. This project will develop an automated, on-line never event prevention system (NEPS) that will interlock the radiotherapy machine to prevent treatment if the patient is not correctly aligned or if the wrong patient plan is loaded, reducing never events by an order of magnitude and directly addressing the AHRQ priority of improving patient safety. Additionally, this project will retrospectively measure the never event rate at UCLA and Veteran's Health Administration (VHA) radiotherapy clinics, testing the hypothesis that radiotherapy never events are significantly under-reported.

Methods

Aim 1: Development of Algorithms

We developed never-event detection algorithms for planar x-ray images from two x-ray setup imaging technologies commonly used in the United States: BrainLAB ExacTrac and Varian OBI. Furthermore, we developed a new algorithm for never-event detection using Cone Beam Computed Tomography (CBCT) setup imaging.

BrainLAB ExacTrac

BrainLAB ExacTrac is used for high-precision stereotactic body radiotherapy, most commonly of the brain and spinal cord. We developed two separate algorithms: one to detect off-by-one vertebral body misalignments (a particularly worrisome error mode) and one general-purpose algorithm to detect misalignments of greater than or equal to 1 cm in any direction to be applied to images of any part of the body.

Convolutional neural net hyper-parameters (input size, number of layers, convolutional stride, learning rate, optimizer, etc.) were defined using experiments on UCLA data. Subsequently, we validated our off-by-one vertebral body algorithm using pooled data from six institutions, including our own. Validation was performed using a "leave-one-institution-out" framework in which the algorithm was trained on data from 5 institutions, and then applied to the 6^{th} institution; the left-out institution was rotated such that all data was used for testing. This approach was used in order to estimate how the model would perform on unseen data from new institutions, ahead of a hypothetical widespread clinical implementation. Algorithm performance was characterized using Receiver Operating Characteristic (ROC) Area Under the Curve (AUC), as well as error-detection sensitivity evaluated with specificity held at 99% and at 95%.

Varian OBI and MV

We developed a CNN-based algorithm to detect off-by-one vertebral body misalignments using Varian OBI planar x-ray images and Varian megavoltage (MV) portal images. At our institution, OBI setup imaging and MV portal setup imaging are predominantly used for alignment of palliative radiation in the abdomen and thorax. Off-by-one vertebral body misalignments are the anticipated error mode. Our algorithms were developed using all OBI and MV portal images containing the vertebral column at UCLA from 2011 to 2021.

Cone-Beam Computed Tomography

Cone-beam computed tomography (CBCT) is used for most radiotherapy setups at UCLA due to superior soft-tissue resolution. Thus, this technology was the focus of much of our work. We developed a general-purpose CNN-based error detection algorithm detecting vertebral body misalignments for images containing the abdomino-thorax and misalignments of 1 cm or more in the head and neck, pelvis, and extremities. This algorithm was developed using image data from UCLA and from Virginia Commonwealth University (VCU). Algorithm performance was characterized using AUC and sensitivity under requirement of 99% specificity.

Our algorithms are based on convolutional neural nets. Our original approach had been based on traditional machine learning using hand-curated features. Around the time this project started, it became clear that deep learning using convolutional neural nets would be the most powerful approach for all computer vision problems. Our work adapting CNNs led as an offshoot to a novel radiotherapy image segmentation algorithm that incorporated a combination of human input and CNNs[1].

Aim 2: Use of Algorithms to find unreported errors in clinical image databases

We searched through retrospective image databases at UCLA and VCU using our CBCT algorithm. 3 years of data was searched at UCLA, and 1 year at VCU. The algorithm was operated such that sensitivity was expected to be greater than 90%, while specificity was at least 95%. All images flagged by the algorithm were reviewed by human experts to determine if a never-event was present. Our original aim was to perform a search for un-reported never-events in image databases at the Hunter Holmes Maguire VA Medical Center in Richomond, VA. Unfortunately, we were not able to secure institutional approval to perform such a search. As an alternative strategy, we performed our search at UCLA and at Virginia Commonwealth University.

Aim 3: Pilot clinical implementation of error detection algorithms

Two channels of clinical implementation are envisioned by our group: first, a real-time implementation at the radiotherapy console which would prevent the radiation beam from being turned on if the expected and actual patient positions did not match. Second, a near-real time implementation that exploits the fractionated nature of most radiotherapy treatments in which treatments are delivered spread over as few as 3 or as many as 20 or more daily fractions. Thus, if treatment errors are detected after a single fraction, more significant systematic errors can be avoided.

In preparation for the expected study of the impact of our algorithm on the clinical treatment workflow, we performed two studies of clinical impact of automation on radiotherapy treatments using implementation science principles. In the first study, we surveyed medical dosimetrists to elucidate barriers and facilities to radiotherapy treatment planning automation[2]. In the second study, we developed and clinically implemented a deep-learning based treatment planning automation program, and studied the impact of the implementation on the clinical workflow[3].

Results

Aim 1: Development of Algorithms

Algorithm performances are shown in **Table 1**. We have completed algorithms for all major x-ray based setup imaging technologies, covering >90% of image-guided treatments in the United States.

Algorithm	Data Source	# Images	AUC	Sens at 99% spec.	Sens at 95% spec.	Publication
ExacTrac spine off-by-one	Multi-institutional	1592	0.985	77.6%	92.8%	[4,5]
ExacTrac >= 1 cm shift	UCLA	27,278	0.997	95.2%	NA	In preparation
Varian OBI vertebral misalignment	UCLA	2544	0.99	NA	99%	Submitted
Varian MV Portal vertebral misalingment	UCLA	1669	0.92	NA	51%	Submitted
CBCT	Multi-institutional	2376	0.995	95%	NA	[6,7]

Aim 2: Use of Algorithms to find unreported errors in clinical image databases

At UCLA, we identified all 3 previously reported never-events, and 4 previously un-reported never-event. No never-events were identified at VCU. The overall alignment never-event error rate at UCLA for CBCT-guided radiotherapy (expressed as a per-radiotherapy fraction rate) was 0.04%+-0.02%. This result was submitted as a manuscript to the International Journal of Radiation*Oncology*Biology*Physics and is currently in revisions.

We analyzed all Varian OBI images containing the vertebral column from treatments from 2011-2021. We identified a total of 1 off-by-one vertebral body misalignments, which was previously un-reported. The vertebral body misalignment error rate at UCLA over this period was determined to be 0.044% (95% confidence interval: [0.0022%, 0.21%]). This result has been submitted as a manuscript to Biomedical Physics and Imaging Express and is currently in revisions.

We are currently analyzing 5-year data from BrainLAB ExacTrac at UCLA using our algorithm. Results have been obtained from 2018-2020 and indicate 2 unreported near-miss wrong-patient treatments, in which a patient was imaged in preparation for delivering a radiotherapy plan intended for a different patient. Investigation determined that in both cases, the involved patients had the same first name and appeared adjacently in the patient schedule. Investigation showed at radiotherapy technologists noticed the discrepancies and halted the incorrect treatments before therapeutic radiation doses were delivered. Complete analysis of the 5-year data is anticipated to be completed by Q1 of 2024 and a manuscript submitted by Q2 2024.

Aim 3: Pilot clinical implementation of error detection algorithms

We developed a software interface that queries the setup image database nightly, locates all setup images acquired in the previous day, applies our algorithm, and prints a web-viewable report. A prototype clinical implementation was performed and the subject of a published manuscript[8]. **Figure 1** shows the workflow of the prototype tool. In preparation for eventual productization of this technique, we performed a series of structured interviews with 19 clinical medical physicists at 16 institutions to elucidate user interface features that would optimize the user experience, and identify barriers and facilitators to clinical use. The manuscript describing these results was submitted to the Journal of Applied Clinical Medical Physics and is currently in revisions. **Figure 2** shows a frequency analysis of the ten most commonly used codes in our thematic analysis of these interviews. **Figure 3** shows barriers and facilitators to the use of automation tools in clinical radiotherapy planning elicited by our survey study[2]; these inform our efforts to commercialize our error detection tool.

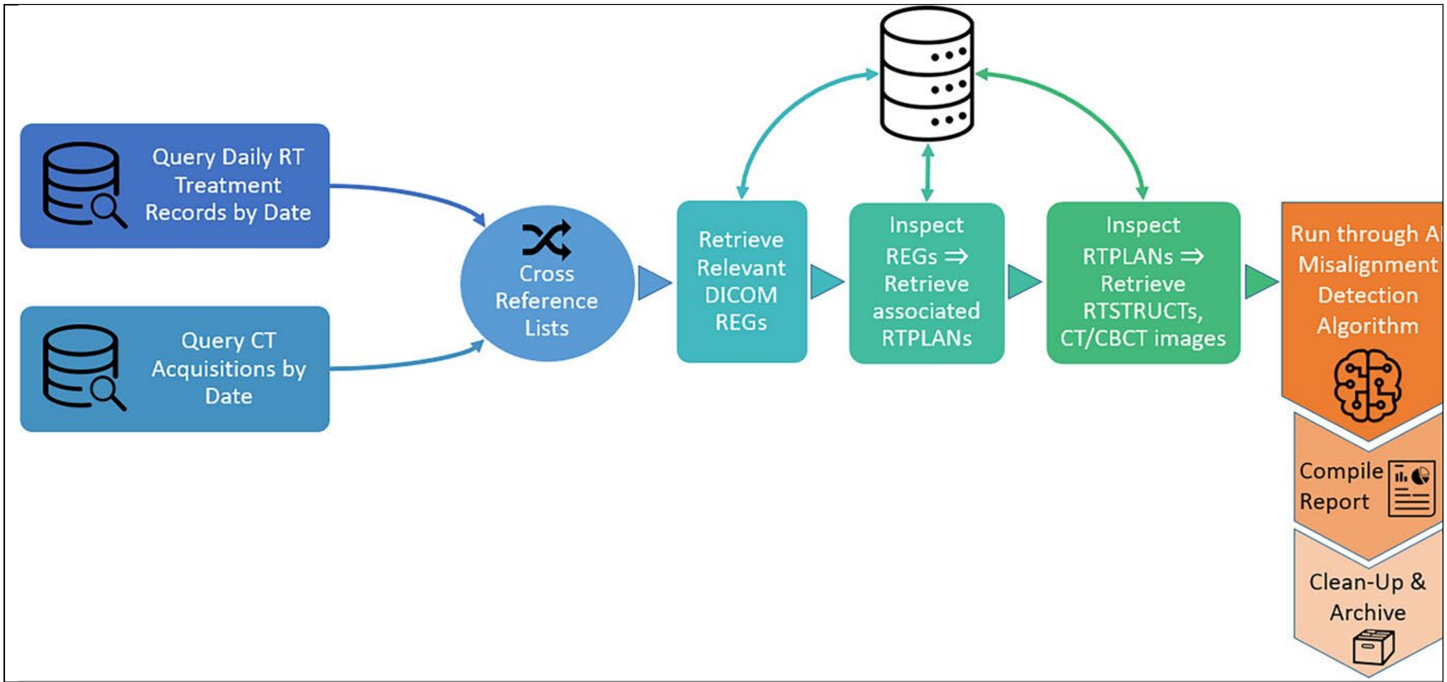

Figure 1: Workflow illustration of the prototype clinical software tool.

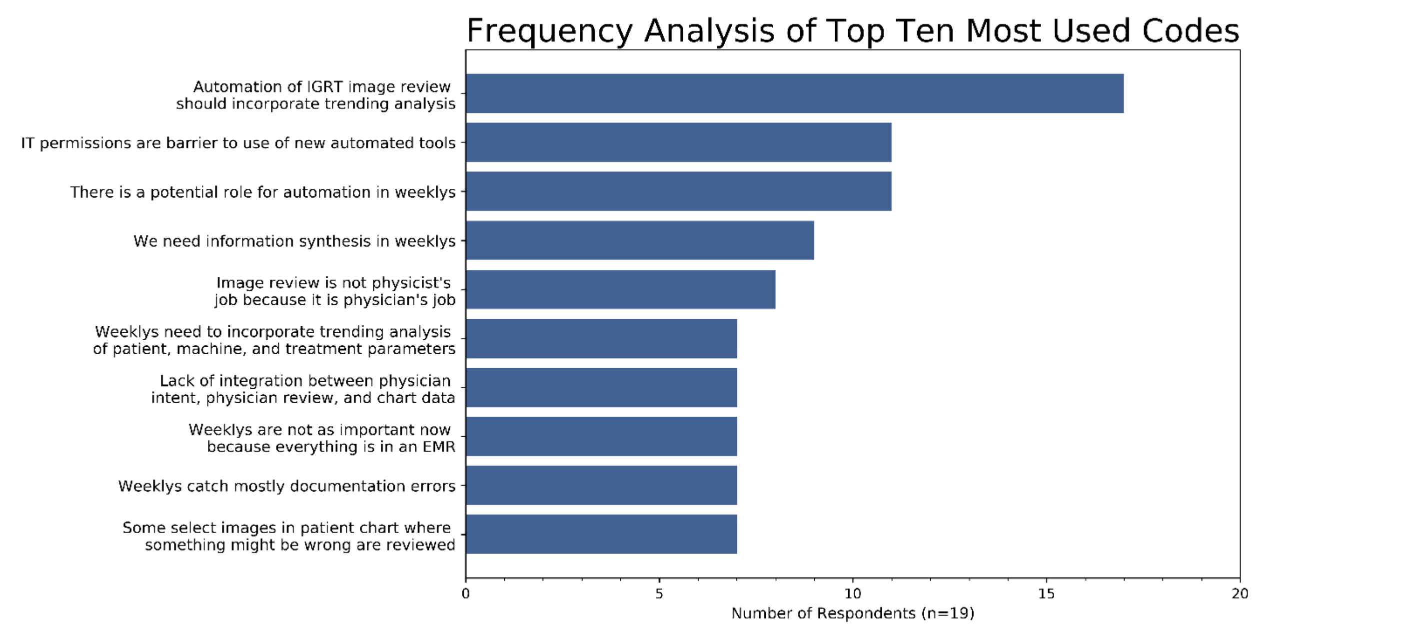

Figure 2: A frequency analysis of the ten most commonly used codes in the thematic analysis.

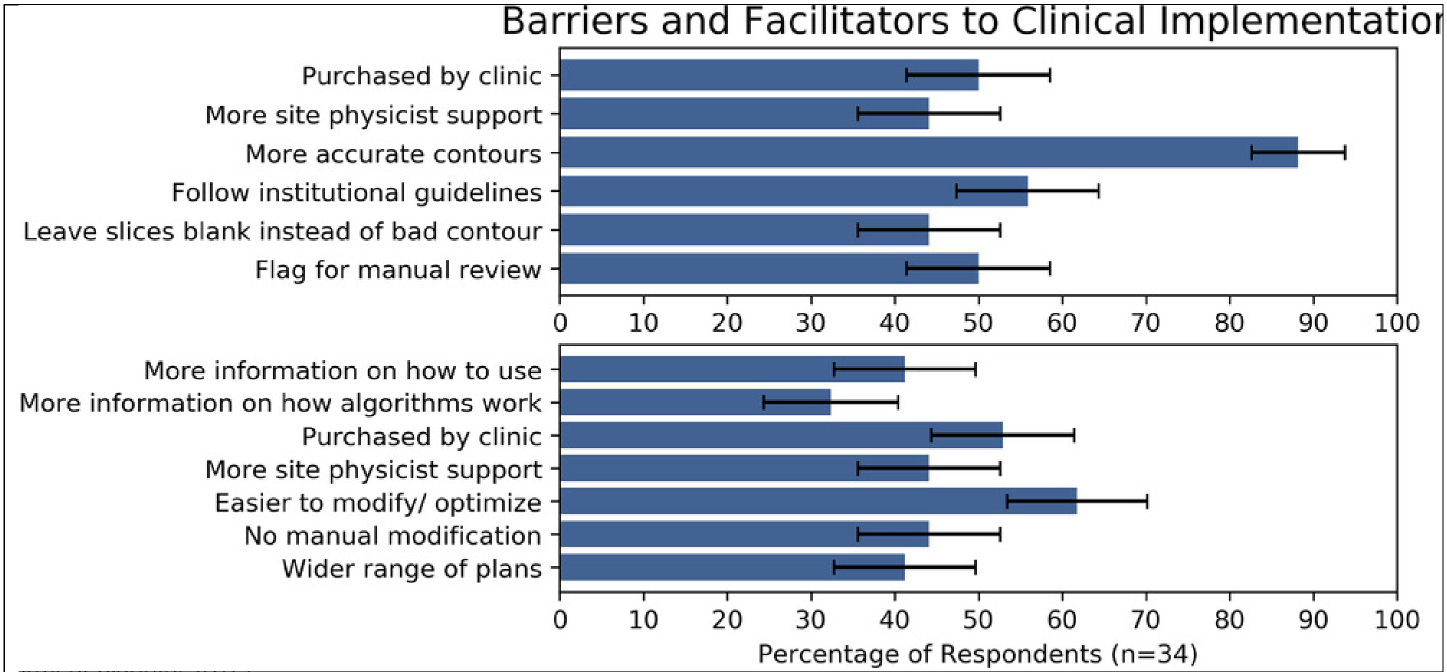

Figure 3: Reported barriers and facilitators to use of automation tools in clinical radiotherapy planning. Error bars represent one standard error.

Real-time clinical implementation has been delayed by a lack of vendor support for making the beam-hold interface available for programming by our team. The beam-hold, or EXGI (external gating) interface is supported by Varian for 3rd party systems to hold treatment beam. However, our team needs documentation describing the abstract programming interface of the EXGI input. Varian agreed to provide such documentation, but has not yet followed through despite repeated inquiries. Thus, our development of a real-time gating interface is currently on hold, however, we do expect to be able to complete it in 2024.

Manuscripts Submitted for Publication

In addition to published manuscripts (listed below), three manuscripts have been submitted and provisionally accepted for publication:

1. Luximon DC, Neylon J, Ritter T, Agazaryan N, Hegde J, Steinberg M, Low D, Lamb J. Results of an AI-Based Image Review System to Detect Patient Setup Errors in a Multi-Institutional Database of CBCT-Guided Radiotherapy Treatments. In revisions at International Journal of Radiation Oncology*Biology*Physics.

2. Petragallo R, Luximon D, Neylon J, Bardach N, Ritter T, Lamb J. Clinical Physicists' Perceptions of Weekly Chart Checks and the Potential Role for Automated Image Review Assessed by Structured Interviews. In revisions at Journal of Applied Clinical Medical Physics.

3. Charters J, Luximon D, Petragallo R, Neylon J, Low D, Lamb J. Automated detection of vertebral body misalignments in orthogonal kV and MV guided radiotherapy: application to a comprehensive retrospective dataset. In revisions at Biomedical Physics and Engineering Express.

Discussion

After successful tests of our algorithm[5], we approached the vendor (BrainLAB) about clinical implementation. The vendor expressed interest in clinical implementation but ultimately did not choose to implement our algorithms or any related techniques; we believe this is due to lack of commercial pressure to implement safety-related product features. Commercialization is the most likely avenue for widespread clinical implementation of our techniques. We have found a potential commercial partner in MIM Software

(Beachwood, OH), a company that provides medical image analysis processing tools, mainly directed at treatment planning and clinical decision support. We are working with MIM on an academic-industrial partnership R01 grant towards commercialization of this prototype tool.

It has been a setback that we have been unable to find a commercialization partner in one of the major radiotherapy equipment vendors (Varian, BrainLAB), despite repeated attempts and despite having made a deliberate choice to put our algorithms in the public domain, thus eliminating intellectual property cost barriers to implementation. Likewise, limited support/documentation of software interfaces that in principle would allow interfacing our algorithms directly to the treatment machines had led to setbacks in clinical prototype implementation in that context.

The result of our retrospective error search showed that errors were underreported: the actual error rate was 66% greater than the reported error rate. However, the absolute error rate remains small, at a population-weighted rate of 0.025% per delivered radiation fraction, or approximately 0.25% per patient depending on fractionation. The impact of an individual error depends on the number of radiotherapy fractions prescribed (which may range from a single fraction, to over 30 fractions), the magnitude of the misalignment, and whether the error was systematic or confined to a single fraction. Our results demonstrate an overall high level of safety in the context of radiotherapy performed in academic institutions.

Conclusions

In conclusion, this project has largely achieved its stated aims, and can be considered a success. The most significant results include a unique absolute measurement of the rate of radiotherapy misalignment errors (all previous measurements having been restricted to reported errors), and the development of a prototype clinical tool to report errors in near-real time, allowing them to be remedied and minimize patient impact based on the fractionated nature of radiotherapy treatments.

Significance

Radiotherapy mis-administrations (never-events) continue to occur in both the planning and delivery of radiotherapy, resulting in patient harm. Our results are the first absolute (i.e. not reliant on voluntary reporting) measurement of the delivery never-event rate, critically informing efforts to improve overall patient safety in radiotherapy and the allocation of finite quality assurance resources. Our error-detection algorithms are novel and unique (no other research or commercial group has produced algorithms that do the same thing) and offer significant opportunity to improve patient safety and/or reduce quality assurance human resources devoted to prevention of radiotherapy delivery never-events.

Implications

The implications of our results are that:

a) Image-guided radiotherapy is in general a highly safe modality with misalignment errors (never-events) at the one-per-1000-patients level.

b) The automated error detection algorithms developed in this grant could be used to redirect some of the human effort devoted towards prevention of misalignment never-events, towards higher value quality assurance tasks.

c) Further effort is needed to commercialize our algorithms in order to maximize clinical uptake.

List of Publications and Products

Published Manuscripts

1. Luximon DC, Abdulkadir Y, Chow PE, Morris ED, Lamb JM. Machine-assisted interpolation algorithm for semi-automated segmentation of highly deformable organs. *Medical physics.* 2022;49(1):41-51.
2. Petragallo R, Bardach N, Ramirez E, Lamb JM. Barriers and facilitators to clinical implementation of radiotherapy treatment planning automation: A survey study of medical dosimetrists. *Journal of Applied Clinical Medical Physics.* 2022;23(5):e13568.
3. Abdulkadir Y, Luximon D, Morris E, et al. Human factors in the clinical implementation of deep learning-based automated contouring of pelvic organs at risk for MRI-guided radiotherapy. *Medical physics.* 2023;50(10):5969-5977.
4. Charters JA, Bertram P, Lamb JM. Offline generator for digitally reconstructed radiographs of a commercial stereoscopic radiotherapy image-guidance system. *Journal of Applied Clinical Medical Physics.* 2022;23(3):e13492.
5. Petragallo R, Bertram P, Halvorsen P, et al. Development and multi-institutional validation of a convolutional neural network to detect vertebral body mis-alignments in 2D x-ray setup images. *Medical Physics.* 2023;50(5):2662-2671.
6. Luximon DC, Ritter T, Fields E, et al. Development and interinstitutional validation of an automatic vertebral-body misalignment error detector for cone-beam CT-guided radiotherapy. *Medical Physics.* 2022;49(10):6410-6423.
7. Luximon DC, Neylon J, Lamb JM. Feasibility of a deep-learning based anatomical region labeling tool for Cone-Beam Computed Tomography scans in radiotherapy. *Physics and Imaging in Radiation Oncology.* 2023;25:100427.
8. Neylon J, Luximon DC, Ritter T, Lamb JM. Proof-of-concept study of artificial intelligence-assisted review of CBCT image guidance. *Journal of Applied Clinical Medical Physics.* 2023:e14016.