Author Manuscript

HHS Public Access
Author manuscript
INFORMS J Appl Anal. Author manuscript; available in PMC 2022 July 15.

Published in final edited form as:
INFORMS J Appl Anal. 2022 ; 52(1): 69–89. doi:10.1287/inte.2021.1095.

Automated and Clinically Optimal Treatment Planning for Cancer Radiotherapy

Masoud Zarepisheh[a], **Linda Hong**[a], **Ying Zhou**[a], **Qijie Huang**[a], **Jie Yang**[a], **Gourav Jhanwar**[a], **Hai D. Pham**[a], **Pinar Dursun**[a], **Pengpeng Zhang**[a], **Margie A. Hunt**[a], **Gig S. Mageras**[a], **Jonathan T. Yang**[b], **Yoshiya Yamada**[b], **Joseph O. Deasy**[a]

[a]Departments of Medical Physics, Memorial Sloan Kettering Cancer Center, New York

[b]Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York

Abstract

Each year, approximately 18 million new cancer cases are diagnosed worldwide, and about half must be treated with radiotherapy. A successful treatment requires treatment planning with the customization of penetrating radiation beams to sterilize cancerous cells without harming nearby normal organs and tissues. This process currently involves extensive manual tuning of parameters by an expert planner, making it a time-consuming and labor-intensive process, with quality and immediacy of critical care dependent on the planner's expertise. To improve the speed, quality, and availability of this highly specialized care, Memorial Sloan Kettering Cancer Center developed and applied advanced optimization tools to this problem (e.g., using hierarchical constrained optimization, convex approximations, and Lagrangian methods). This resulted in both a greatly improved radiotherapy treatment planning process and the generation of reliable and consistent high-quality plans that reflect clinical priorities. These improved techniques have been the foundation of high-quality treatments and have positively impacted over 4,000 patients to date, including numerous patients in severe pain and in urgent need of treatment who might have otherwise required longer hospital stays or undergone unnecessary surgery to control the progression of their disease. We expect that the wide distribution of the system we developed will ultimately impact patient care more broadly, including in resource-constrained countries.

Keywords

radiotherapy cancer treatment planning; intensity modulated radiation therapy; hierarchical optimization; multicriteria optimization; large-scale optimization; mixed-integer nonlinear programming; Edelman Award

Memorial Sloan Kettering Cancer Center

Memorial Sloan Kettering Cancer Center (MSK) is a geographically diverse patient-care network consisting of a main campus in Manhattan (in New York City) and 20 facilities in New York and New Jersey. MSK is the largest and oldest private cancer center in

zarepism@mskcc.org (MZ).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

the world and is regarded as one of the world's leading cancer centers with a tri-fold mission of patient care, research, and training. Each year, MSK treats more than 25,000 patients, including patients suffering from all types of cancer (approximately 400) and trains over 3,000 medical students, residents, postdoctoral researchers, and PhD and MD-PhD candidates. MSK annually treats about 11,000 new patients with radiotherapy.

What is Radiotherapy Cancer Treatment?

Radiation therapy (radiotherapy), which uses radiation to sterilize cancer cells by damaging their DNA, is one of the major modalities of cancer treatment. Radiotherapy can be prescribed as the main treatment or can be given in combination with surgery and/or chemotherapy, depending on the details of the disease and the overall health of the patient. Radiotherapy is often preferred when surgery is unlikely to be well-tolerated. It can also be given prior to surgery to reduce the size of the tumor to be resected (i.e., cut out or partially cut out), thereby making the surgery safer. Another emerging use is to ablate (i.e., destroy) isolated metastatic tumors that have spread from the original tumor. In addition, radiotherapy is often given to provide palliative (i.e., noncurative) care for patients, using reduced-dose treatments with the goal of decreasing pain or the impact of disease on normal functions, thereby improving the quality and dignity of the end-of-life process.

Radiotherapy uses precisely shaped, high-energy beams to kill cancer cells, and is usually given from a machine (medical accelerator) outside the body (external-beam radiation therapy), most commonly by photons (i.e., X-rays). External-beam radiation therapy is the most common form of radiotherapy treatment. Intensity-modulated radiation therapy (IMRT), which is the most commonly used form of external-beam radiation therapy, is implemented using advanced computer programs to design and deliver converging radiation beams from different angles using complex patterns of radiation intensity.

Computerized radiotherapy planning allows the delivery of effective doses of radiation while intricately shaping the dose distribution to each patient's unique tumor anatomy, leading to limited morbidity in nearby healthy tissues and organs. Multileaf collimators (MLCs) are used to control the shape and intensity (amount per unit time of exposure) of the radiation coming from each beam that the medical accelerator delivers (see Figure 1). This highly complex pattern of shape and radiation intensity is determined during the planning process and is unique for each patient. The MLC is mounted on the head of a rotating linear accelerator gantry and uses a set of metal fingers or "leaves" (which are thick in the beam direction but thin in cross section) that can move laterally in and out. Analogous to how a window shade blocks or emits sunlight depending on the shade's position, a leaf will block or emit radiation transmission depending on the leaf's position. Such leaf motions, with the radiation beam on, thereby build up a pattern of delivered, modulated radiation to the patient. Typically, the gantry rotates around the patient and makes multiple stops at predetermined angles (typically 7–11) to deliver an angle-specific modulated pattern of radiation via moving the leaves while the radiation beam is 'on' (this technique is known as sliding window). An expert treatment planner determines the beam angles according to the patient's geometry and location of the tumor and surrounding organs and subsequently algorithms optimize the leaf movements for all the angles together. The total radiation

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

delivery time varies significantly depending on many factors, including prescription dose and machine capability, but it is typically less than 10 minutes.

MSK helped pioneer the development of IMRT and was one of the first institutions to clinically deploy IMRT delivery on conventional linear accelerators. It has used IMRT to treat prostate, breast, and other cancers since the mid-1990s (Leibel et al. 2002).

The radiotherapy process comprises four main steps, as we illustrate in Figure 2.

1. Imaging (Figure 2A): Patient imaging is performed, typically using computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET) imaging, or a combination of these techniques, to obtain a detailed digital representation of the patient's body and the location and extent of the disease. For ease of explanation, we will write this paper as if the disease has one contiguous location (one tumor) although the work applies to multiple-tumor treatment.

2. Contouring (Figure 2B): The tumor and surrounding healthy tissue are accurately localized by drawing their borders (contours) on the digital image.

3. Treatment planning (Figure 2C): This step customizes the delivery-machine parameters (radiation beams angles, MLC leaf movements) for each patient. The leaf movements include the speed of the leaves and thus determine the duration of radiation delivery. An expert treatment planner customizes the machine settings using FDA-approved treatment planning software, and later consults with a physician for final approval or instructions to change the plan. This step, which is the focus of this project, is time consuming and labor intensive, and plan quality is heavily dependent on the planner's experience and skill.

4. Treatment delivery (Figure 2D): To deliver the treatment plan, the radiation delivery machine rotates around the patient and delivers radiation with different MLC leaf positions and shapes from different directions. Treatments are typically fractionated, ranging anywhere from a single session in which a high-dose is delivered, to many weeks of more modest dose delivery per session.

Radiotherapy Treatment Planning

The main objective of radiotherapy is to deliver a sterilizing dose of radiation to the tumor to kill cancerous cells without harming the surrounding healthy organs and tissues. Each cancer patient's tumor is unique in terms of its shape and location and the surrounding healthy tissue. The machine settings, including the beam angles and beam shapes (defined by MLC leaf movements), must be customized based on the patient's image data and the physician's dose prescription for both the tumor and nearby healthy normal tissue. This process, referred to as *radiotherapy treatment planning*, is performed by a *medical physicist or dosimetrist (hereafter referred to as the "planner")*, a highly trained medical professional with a physics and medical educational background. This process is a crucial step impacting both the treatment outcome and the patient's quality of life.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Radiotherapy Treatment Plan Metrics

For a given set of machine settings (accelerator and MLC parameters), the delivered radiation dose is simulated using a FDA-approved treatment planning system and is usually presented to the user (planner) in terms of dose distribution (Figure 3A) and dose volume histograms (DVH) (Figure 3B) for evaluation.

The dose distribution is shown as a three-dimensional (3D) color map laid on top of the digital image, representing in a color-coded fashion the amount of radiation dose delivered to each part of the patient's body (Figure 3A). Ideally, we want all parts of the tumor to receive the prescribed radiation dose (40 Gy for the example in Figure 3). However, because of the need to avoid nearby normal tissue, some loss of uniformity is unavoidable. Instead, some parts of the tumor receive more or less than the prescribed radiation dose, within an acceptable deviation. We want nearby healthy tissue to receive as little radiation as possible. However, due to the steady attenuation of radiation beams as they traverse the body, this is not possible. The optimization problem is to achieve the needed tumor dose while minimizing the dose given to healthy tissue. The problem is further complicated because different tumors require more or less radiation for control, and healthy tissue have different inherent tolerances to radiation.

The planner can visually inspect the dose distribution by scrolling through different images. A DVH is a convenient two-dimensional (2D) plot for plan evaluation and can be easily calculated from the 3D delivered dose. As Figure 3B shows, a DVH curve is associated with each organ or tissue specifying the fraction of the volume of the organ (y-axis) that receives more than the dose indicated on the x-axis. For example, for the left lung, the point indicated by the star shows that 10% of the left lung receives at least a 20 Gy radiation dose (V(10%) = 20Gy). It is important that the DVH curve for each organ is below the physician's prescribed DVH points of that organ; in this example, for the left lung, the DVH curve should be below the point indicated by the star, which means that at most 10% of the left lung should receive a dose of 20 Gy or more. For tumor target volume, the point indicated by the star shows that 90% of the tumor receives at least 40 Gy (the prescribed dose in this example). Given that the tumor should receive its target volume, the curve should be above that prescribed DVH point.

Current Clinical Practice and the Need for the New Paradigm

In modern radiotherapy machines, MLCs usually have 60 leaf pairs with the leaf width as small as 0.25 cm and as large as 1.0 cm projected to the patient distance. Each leaf can be positioned with sub-millimeter accuracy. In the 1980s, Anders Brahme and others recognized that high-dose regions could be shaped more closely to the tumor target if modulated beams of radiation were designed carefully using computer algorithms (Brahme et al. 1982). This led to the development of IMRT and the associated treatment planning process in which a mathematical model and optimization algorithm determine the machine parameters (including MLC leaf positions and movements) to best achieve the specified clinical goals.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

From the 1990s until today, the most common IMRT planning optimization technique, which is still used in most current FDA-approved treatment planning systems, is the *weighted-sum* formulation, wherein the algorithm optimizes the machine parameters that minimize the deviation of the realized dose distribution beyond the physician's prescribed clinical goals (i.e., the weighted summation of the deviations of all the specified goals). Although IMRT planning has represented a paradigm shift in radiotherapy, and working with the optimization parameters (i.e., clinical goals and their weights) is more intuitive than working with the machine parameters, it still involves a tremendous amount of trial-and-error work and the manual tuning of the goals and weights. Accordingly, the treatment planner defines a set of goals and weights and runs the optimization algorithm. Then, the resultant dose distribution and DVH are generated for evaluation. If they are not satisfactory, the planner typically tweaks the parameters and runs the optimization problem again until a satisfactory plan is generated, or the process is truncated due to time constraints. For example, if the mean dose to the lung exceeds the physician's tolerance value, then the planner might increase the weight associated with the lung mean dose to reduce it, despite the possibility that doing so might introduce undesirable issues with the plan (e.g., an increase to the esophagus mean dose).

The current treatment planning process thus requires extensive iterative manual tuning of the parameters to generate an acceptable plan, and often requires hours of active trial-and-error experimentation by the planner. Consequently, the quality of treatment plans can vary extensively among planners because the process is heavily dependent on the planner's skill and experience. Thus, the entire planning process (which includes other steps not described here) is time consuming and labor intensive, and typically requires multiple days. The resulting delays can prolong patient hospitalization, or even necessitate surgery to control tumor progression or pain for patients in need of urgent treatment.

For the foregoing reasons, the main objective of this project was to develop a fast, inexpensive, high-quality, and unbiased treatment planning system, as we outline below.

1. Fast: Reduce the time needed to generate a high-quality plan, particularly for patients in severe pain and in urgent need of treatment.

2. Inexpensive: Reduce healthcare costs and resource usage.

3. High quality: Maximize the clinical outcome by delivering sufficient radiation to the tumor and minimize the treatment side effects (e.g., swallowing problems, rectum bleeding, lung damage) by reducing unintended radiation to the surrounding healthy organs. This should be accomplished in a mathematically inarguable way, so that planners and physicians have confidence that the solution is the best possible.

4. Unbiased: Reduce the variability of treatment quality across clinics and among planners within each clinic.

Our method is based on hierarchical constrained optimization. We internally refer to this method and our project as expedited constrained hierarchical optimization (ECHO). Other methods that have been developed to speed up treatment planning include (1)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

multiple-criteria optimization (MCO) (Craft and Bortfeld 2008, Monz et al. 2008), and (2) knowledge-based planning (KBP) (Appenzoller et al. 2012, Ge and Wu 2019). MCO generates a pool of Pareto optimal plans based on a small number of key clinical objectives that require a trade-off. A user then navigates among the Pareto plans, selecting a preferred plan. Although MCO improves the planning process, the planner still must do the planning and the plan quality is therefore dependent on that planner's experience. KBP, in contrast, uses artificial intelligence (AI) and statistical methods to improve the radiotherapy planning process. In the AI-based solution, an expert system learns characteristics of previously accepted plans to set planning parameters. Such a solution has been shown to increase efficiency and reliability above the current clinical practice (i.e., the trial-and-error, open-loop formulation). The AI-based solution is computationally faster than ECHO, because it does not require solving large-scale constrained optimization problems. However, important drawbacks of the AI-based approach are that it does not guarantee clinical optimality for individual patients, and it relies heavily on the quality of a reference dataset, which it seeks to mimic. More importantly, ECHO is easily adaptable to dynamic and changing clinical practices in terms of prescription doses and healthy organ criteria; in an AI-based system, such changes would require retraining of the entire system.

History of Operations Research in Radiation Therapy

Given that so many journal and publication resources are available today, accurately tracking the developments in any field is difficult. We provide a brief history developments of operations research (OR) in radiotherapy, which is accurate to the best of our knowledge. The use of OR in radiotherapy predates the advent of IMRT. Figure 4 provides an estimate of the number of OR publications on radiotherapy. The search was performed on the Scopus website (www.scopus.com) using the following terms:

> TITLE-ABSTRACT-KEYWORD ((“radiation therapy” OR “radiotherapy”) AND (“optimization” OR “operations research”)).

We highlight some important publications and events in Figure 4. In the 1980s, IMRT was developed (Brahme et al., 1982) and the medical physics community applied OR to optimize IMRT (Webb 1989). In the 1990s, members of the OR community became aware of the application and their interest gained momentum after the *Siam Review* published Shepard et al. (1999) and NCI/NSF held the workshop discussed in Lee et al. (2003a). Up to 2020, *Siam Review* journals included more than 40 publications about radiotherapy and INFORMS journals included more than 60 articles. For detailed information, we refer the reader to review papers published on this topic (Bortfeld 2006, Ehrgott et al. 2010, Romeijn and Dempsey 2008).

Our project has benefited significantly from the previous developments in the field. As a result of this work, new ideas and OR techniques have been introduced, and these techniques have been implemented clinically so that patients can benefit from advanced OR tools. In the Patient Safety and FDA Considerations section, we review the OR tools developed for this project and in the Summary of Scientific Contributions section, we summarize the scientific contributions of this project. Most of the scientific contributions have been published in peer-reviewed journals and we refer the reader to thefor detailed information (Dursun et al. 2021, Hong et al. 2019, Mukherjee et al. 2020, Taasti et al. 2020a, Taasti et al. 2020b, Zarepisheh et al. 2019b). Zarepisheh et al. (2019a) describe a published patent application.

Automated Treatment Planning: Challenges and OR Solutions

Table 1 summarizes our six biggest challenges and the OR solutions we developed before implementing ECHO in the clinic. We have clinically implemented solutions to the first three challenges and are using them to treat patients in our daily clinical routine. We review them in this section. The second set of three challenges are in the research and development process; we review them in Appendix B. The details can be found in the peer-reviewed journals cited in the last column of Table 1.

Challenge 1: Conflicting Objectives

Challenge: As radiation travels through the body, it is delivered to the tumor and also to nearby tissue through which it passes. Therefore, there is an intrinsic conflict between the goals of delivering enough radiation to the tumor and minimizing the unintended radiation to healthy organs. An optimal plan must achieve clinically meaningful trade-offs between these conflicts to cure the tumor without unacceptable side effects to the healthy organs.

Solution: This conflict is formulated as a hierarchical constrained optimization problem (internally ECHO). For each patient, the machine parameters are optimized by solving two large-scale sequential constrained optimization problems. Figure 5 provides a high-level description of the optimization problems. The first-level optimization problem (Step 1) guarantees an adequate radiation dose to the tumor; the second-level optimization problem (Step 2) minimizes high radiation doses to critical healthy organs. Some tissue-tolerance dose values are considered as medically so important that they are implemented in the optimization problem as firm constraints (with maximum values). The plan delivery feasibility and efficiency are ensured using a smoothness constraint, which controls the total variation of intensities among the neighboring beamlets. We provide the detailed mathematical formulations of Step 1 and Step 2 in ECHO Mathematical Formulation in Appendix A.

Novelty: Hierarchical constrained optimization (also called prioritized optimization or lexicographic optimization) is a classic optimization technique used to address many complex multicriteria optimization problems. The applicability of this technique in radiotherapy optimization was introduced in Deasy (2000) and Deasy (2002). It was later expanded to four-step optimization models in Wilkens et al. (2007) and Clark et al. (2008). Their objectives were to (1) maximize tumor coverage, (2) minimize doses to high-priority noncancerous organs, (3) minimize doses to other noncancerous organs, and (4) smooth out the intensity profile (also known as the fluence map). Breedveld et al. (2009) introduce another variation of this technique—the so-called wish-list where a series of optimization problems is solved to meet, as much as possible, the predefined clinical goals, and is followed by solving another series of optimization problems.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Our approach consists of two steps, thus making it computationally more attractive. Furthermore, we clinically implemented, validated, and seamlessly integrated our technique with our existing FDA-approved clinical software.

Challenge 2: Computational Issues and Modeling Inaccuracies

Challenge: The resultant optimization problems are large with hundreds of thousands of variables and constraints (typically approximately 100,000 variables and 500,000 constraints). To ensure a rapid turnaround and timely patient treatment, the problem must be solved in a clinically reasonable time frame (one to two hours). The optimization model also has unavoidable, but well-understood, inaccuracies. For example, the optimal beamlet intensities may not be perfectly reproduced using MLC leaf movements due to mechanical limitations. There are nonnegligible radiation leakage effects through the leaves, which cannot be captured and modeled upfront and will only be realized after calculating the final dose. Such inaccuracies could degrade the result and lead to a suboptimal plan.

Solution: An influence matrix contains the dose delivered to each voxel from each beamlet of unit intensity. The original influence matrix is dense and large with tens of thousands of columns (corresponding to beamlet intensities) and hundreds of thousands of rows (corresponding to tissue voxels), making it computationally expensive and challenging. To address this, we truncate small elements of the matrix to zero to speed up the optimization. However, this introduces inaccuracies in the model. We later compensate for the discrepancies between the optimization results and the final results by solving an equivalent unconstrained optimization problem counterpart generated using Lagrangian multipliers (Figure 6).

After solving Step 1 and Step 2 (nonlinear constrained optimization problems), the optimal beamlet intensities are then imported into an FDA-approved treatment planning system for MLC leaf-movement calculations and accurate final-dose calculations. Then, we calculate the discrepancy between the final dose and the optimization dose, solve a correction step (Step C) optimization problem, and import the optimal beamlet intensities into the FDA-approved treatment planning system for final-dose calculations (Figure 6). In ECHO Mathematical Formulation in Appendix A, we provide the detailed mathematical formulation.

Novelty: The dose-discrepancy issue is well-known as part of the conventional weighted-sum method and is usually handled by periodically performing a full-dose calculation during the iterative optimization process and incorporating the discrepancy into the optimization (Siebers et al. 2002). However, this is not consistent with the constrained hierarchical optimization approach because updating the dose values could invalidate the constraints added in the previous steps and thus make the problem infeasible. To the best of our knowledge, using Lagrangian multipliers to address this issue is a novel solution.

Challenge 3: Nonconvexity of Objective Functions and Constraints

Challenge: Some plan-quality metrics and constraints, that is, DVH constraints (e.g., no more than x% of organ Y shall receive more than dose D), are inherently combinatorial

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and nonconvex, resulting in computationally intractable large-scale mixed-integer nonlinear programming (MINLP) problems that cannot be solved in the restrictive clinical timeline.

Solution: We introduce an auxiliary optimization problem, referred to as Step 0, to the two-step optimization model. Step 0 essentially solves Step 1 but supplements it with a convex approximation of the nonconvex DVH constraints (i.e., relaxing the integrality of the binary variables in the corresponding MINLP). Although this convex approximation does not guarantee the satisfaction of the constraints, it provides crucial initial information about which voxels are likely to receive doses of radiation below that particular DVH-constraint dose threshold. Subsequently, maximum-dose constraints are imposed on these low-dose voxels to ensure the fulfillment of the DVH constraints in Steps 1 and 2 (Figure 7).

The proposed algorithm is computationally tractable and provides a solution in the proximity of the MINLP-based ground truth solution, often in less than one hour. For a low-resolution case of a spine tumor, we obtained the ground truth solution by solving the MINLP problem in 15 hours, whereas the proposed algorithm converged in only two minutes with an equivalent solution (Figure 8).

Novelty: The MINLP has been introduced in the past to handle DVH constraints (Langer et al. 1990). Lee et al. (2003b) proposed using a specialized branch-and-bound MINLP solver to improve the performance of the MINLP. Despite these advances, solving the MINLP problem for an actual large-scale clinical-data problem in an acceptable amount of time remained a difficult and unsolved problem. Combining the relaxed MINLP problem with an effective heuristic, by taking advantage of the special characteristics of the problem and the underlying nonconvex constraint, to solve the real clinical problem within the acceptable clinical timeframe is a novel approach. To the best of our knowledge, this is also the first time the performance of a heuristic algorithm to handle the DVH constraint has been verified by comparing it against the ground truth.

Our clinical implementation of ECHO includes all the technical solutions mentioned in Challenges 1–3 and solves Steps 0–2 and the correction step for each patient. All the resultant optimization problems are constrained nonlinear problems with only continuous variables and are solved using AMPL and the interior point algorithm implemented in Knitro.

Patient Safety and FDA Considerations

Any new technique used in a clinical environment should be FDA compliant and safe for patients. To resolve this issue, we have integrated our new algorithm (ECHO) with an FDA-approved commercial treatment planning system named Eclipse (Varian 2021) (Figure 9). We use the application programming interface (API) of Eclipse to extract patient data needed for optimization; we then send the optimal beamlet intensities back into the system after optimization for a final, accurate, dose calculation and plan review. The ultimate plan evaluation and delivery are performed in the FDA-approved planning system; therefore, additional FDA approval to use ECHO is unnecessary.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The Department of Medical Physics at MSK has a long tradition of developing and clinically deploying fit-for-purpose technical processes for radiotherapy and other biomedical engineering challenges, including an early IMRT planning system. It follows software quality assurance (QA) methods and performs extensive predeployment testing to ensure safe usage. In addition, the challenge of building an IMRT planning engine that is dosimetrically accurate is extensive in that many minor aspects must be considered. MSK has a rigorous QA procedure to ensure the safe delivery of radiotherapy to all patients. For example, all plans for patients who will receive their entire radiation treatment in a single session are delivered first on the treatment machine without the patient present using an experimental setup to measure and assess the dose accuracy. After each IMRT treatment fraction, the machine delivery logs (including all the delivery information) are analyzed to ensure that the radiation has been delivered according to the plan.

Benefits

In this section, we review the main benefits provided by ECHO to our patients and to MSK. Table 2 provides a summary of the benefits.

Patient-care Quality: Consistent High-quality Plans

Introducing any new radiotherapy regimen into a clinic requires rigorous preclinical study and testing. For each cancer tumor site (e.g., paraspinal, prostate, lung), before deploying ECHO in our clinic, we run a preclinical study to ensure that ECHO's use improves workflow, saves resources, and improves planning quality. To this end, we randomly identify a set of patients who have received treatment in the past using the previous clinical system (i.e., manually generated plans) and compare their results with automatically generated ECHO plans. We perform this comparison using some common, established clinical metrics to evaluate the radiation dose. As an example, for plans for spinal tumors, as part of our preclinical study, when we compared 75 ECHO plans with manually generated plans that were used for treatment, we found that ECHO typically resulted in better, more consistent plans (Figure 10).

Figure 11 illustrates how ECHO can improve the quality of treatment plans by comparing an ECHO plan with a manually generated plan in terms of dose distribution and the resulting DVHs. The dose distribution demonstrates that an ECHO plan gives less dose to the esophagus. The DVH plots illustrate that ECHO delivers more uniform dosage to the tumor (i.e., dosage closer to the prescription).

Addressing Patient Needs: Expedited Treatment for Patients in Pain

The optimal management of spine tumors is often time sensitive because of the delicate nature of the spinal cord. Even a mild compression of the cord can result in spinal cord dysfunction, and permanent disability can result from if cord compression is not relieved within 24 hours, thus representing an oncologic emergency. Hence, clinical management can be influenced by the time required for a treatment plan. In cases where rapid progression of disease near the spinal cord is a concern, patients who otherwise might not need spinal surgery may undergo surgery if radiotherapy cannot be planned and executed in a timely

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

manner. With the advent of ECHO planning, high-quality plans can be generated rapidly, sparing some patients from spinal surgery. Prior to ECHO planning, some patients were kept in the hospital for additional days to allow the treatment planning process to run its course before starting radiotherapy. Hence, the time savings resulting from ECHO's use have reduced the number of hospital days in such cases, generating meaningful savings in healthcare dollars and a better patient experience. Another advantage of the ECHO process is that treatment plans are of equally high quality independent of the planner, thereby effectively expanding the resource pool of planners. Furthermore, in healthcare networks with multiple facilities, an ECHO plan could be generated at any site within the network, independent of where the patient's images were generated, the planner is located, or at which site in the network the patient will receive treatment. This flexibility allows for an efficient utilization of resources, and no time is lost waiting for an available planner.

Since 2017, ECHO has been used to produce more than 2,800 clinical plans on an expedited schedule for patients with tumors near the spinal cord, enabling quicker commencement of treatment, which positively impacted the patients' quality of life. The ability of ECHO to produce a high-quality optimal plan quickly and reliably enabled some patients to be treated on the same day of their CT image acquisition and further improved the patients' clinical experiences. The ability to provide same-day imaging and treatment has allowed the development of new treatment paradigms, such as using radiation instead of, or in tandem with, surgical intervention. For example, preoperative radiation may offer advantages over postoperative radiation for hip-metastases patients who are at high risk for hip fracture and who require urgent surgical stabilization. Preoperatively, the gross hip tumor is intact, and a more compact radiation treatment volume can be used instead of the postoperative setting, when the tumor area has been surgically violated and more tissue must be irradiated to control any remnants of disease. Patients need not return for postoperative imaging and treatment, because imaging, planning, and treatment were performed in one day, just prior to surgery. This is particularly useful when postoperative patients require rehabilitation after surgery because radiotherapy cannot be administered during the period of postoperative rehabilitation, a process that typically lasts several weeks. Hence, the rapid high-quality treatment planning that ECHO provides has meaningfully impacted patient care, by reducing the number of surgeries performed and length of inpatient hospital stays, and by increasing the flexibility of planning resources.

Financial and Resource Benefits

Prior to implementing ECHO, MSK, like many cancer centers, used a commercial planning solution that required the manual tuning of the optimization weights of different tissue volumes in the planning problem. Because the clinical trade-offs between different types of healthy tissue can be patient specific, extensive iterative manual tuning is required to generate a satisfactory plan. Using ECHO, all the critical clinical criteria are enforced as constraints and satisfied without any manual tuning. The desirable criteria, including more dose to tumor and less dose to healthy tissue beyond the specified clinical criteria, are then, as much as possible, optimized sequentially. This results in a satisfactory plan without any manual tuning or iterations.

Author Manuscript

We have studied the impact of ECHO on the efficiency of the planning process, and we conservatively estimate that when ECHO performed the planning, productivity improved by at least 15% per planner. ECHO is currently in a ramp-up phase, accounting for about 15% of all MSK IMRT radiotherapy treatments. We expect to increase this to above 80% of all our IMRT planning (i.e., to 53% of all MSK treatment plans) by the end of 2022, and we conservatively expect that ECHO will save the equivalent effort of at least 13.5 full-time planners at MSK (15% of our planners). The United States has about 4,500 treatment planners and the use of ECHO could potentially slow the required growth in number of planners needed to treat increasing numbers of cancer patients. The worldwide impact of ECHO, particularly in resource-limited regions, would likely be measured as much or more in increased treatment capacity and the ability to better meet the need for radiotherapy to treat cancer. Thus, given the need for radiotherapy treatments for more than six million cancer patients worldwide each year, a number that is projected to increase, ECHO has the potential to impact millions of lives annually.

The reduced workload enabled by ECHO may also be used partly to innovate, enabling the introduction of new processes to improve treatments that are currently too resource intensive, such as automating session-by-session monitoring of accumulated dose quality; most patients receive between 3 and 30 separate fractions (i.e., sessions) of radiotherapy.

Since its inception in 2017, ECHO has automatically generated more than 4,000 high-quality plans in our clinic and the number of ECHO plans is increasing steadily (Figure 12). It is currently being used for five disease sites: spinal, metastatic, prostate, lung, and head-and-neck tumors.

Transportability and Potential Global Impacts

Transportability

ECHO is currently integrated with Eclipse, one of the world's most common commercial treatment planning systems. ECHO could be used in hospitals with this treatment planning system, or it could be integrated with other FDA-approved systems. Physician-driven clinical criteria (i.e., prescriptions) are embedded in ECHO as template files, but they can be customized easily according to practice variations. Some software development is needed to make ECHO compatible with other existing commercial treatment planning systems, although the optimization algorithm would still be applicable. The software adaptations required would be mainly to deal with data communication (i.e., exporting the data for optimization and importing the optimized plan back for final calculation). The premise of ECHO and the algorithm design is independent of the treatment planning system and delivery machine. ECHO can also be adapted for use with other less frequently used types of external-beam radiotherapy (e.g., proton beam, carbon ion, combined magnetic resonance imaging-linear accelerator systems).

ECHO also provides an excellent platform for future technical advances in radiotherapy. Treatments could be monitored offline, after treatment fractions, for necessary planning updates based on a patient's changing anatomy or on any errors that are found in the treatment fractions delivered based on routine imaging at the time of treatment.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Global Impacts: Potential Benefits for Resource-limited Countries/Hospitals

Currently, according to the International Organization of Medical Physics (IOMP), except for North America, Europe, and a few countries in Asia, more than 75% of countries worldwide do not have accredited training programs for medical physicists or dosimetrists. Hence, most regions in the world lack highly skilled professionals to produce clinical plans for radiotherapy patients. Therefore, resource-limited countries can benefit the most from using ECHO to help improve their standard of the care while reducing the required resources.

As noted, physician-provided clinical criteria are embedded in ECHO, and hospitals can either customize them according to their needs or use them as is to transfer and share knowledge and expertise across institutions (e.g., doses to the tumor and dose constraints to healthy tissues). As a global leader in cancer care, MSK has been at the forefront of radiotherapy for decades with vast experience and expertise. Our goal of porting ECHO to other clinical treatment planning systems would carry MSK's excellence of radiotherapy planning to a wide network of users.

Summary of Scientific Contributions

We have developed and clinically implemented a new paradigm for radiotherapy treatment planning, which we refer to as ECHO. ECHO replaces trial-end-error IMRT planning methods with an approach that guarantees high-quality treatment plans that respect clinical priorities. Using ECHO, high-quality plans are generated rapidly and reliably independent of the planner's experience. It enables expedited treatment for patients in severe pain and in urgent need of treatment who otherwise might have to undergo surgery to control their disease progression or stay additional days in the hospital awaiting the start of their radiotherapy treatment. ECHO furthermore reduces the huge variability in treatment quality among human planners within and across clinics. It has impacted over 4,000 patients to date and will be expanded to the great majority of all radiotherapy treatment planning at MSK in the next two years. The ECHO system will ultimately be explored as a way to impact patient care more broadly, possibly with a commercial partner, including in resource-constrained countries where access to highly skilled radiotherapy planners is limited and cost-efficient resource utilization is required to meet cancer treatment needs.

Advanced OR tools are the main building blocks of ECHO. To the best of our knowledge, ECHO is a first of its kind in terms of its clinical implementation and seamless integration with an FDA-approved treatment planning system. The scientific contributions of ECHO fall into three categories:

1. Building on OR tools that have already been applied in radiotherapy:
    - Mathematical modeling of radiation and physics (Zarepisheh et al. 2019b)
    - Hierarchical constrained optimization (Zarepisheh et al. 2019b)
    - Robust optimization (Taasti et al. 2020a)

- Mixed-integer nonlinear programming (Mukherjee et al. 2020)

2. Identifying OR tools never applied in radiotherapy

- Bayesian Optimization (Taasti et al. 2020b)
- Sequential convex programming (Dursun et al. 2021)

3. Pushing the boundaries of OR

- Lagrangian methods to correct modeling inaccuracies (Zarepisheh et al. 2019b)
- Customized SCP (Dursun et al. 2021)
- Heuristic MINLP (MINLP + application domain knowledge) (Mukherjee et al. 2020)

Acknowledgments

This work was partially supported by MSK Cancer Center Support Grant/Core Grant from the NIH (P30 CA008748). The authors thank Aaron Burciaga and Michael Prokle for providing many helpful comments and suggestions as coaches for Edelman 2021 competition.

Appendix A: Mathematical Modeling

General Mathematical Modeling

In this section, we briefly introduce some basic and well-known mathematical modeling of the IMRT problem. In a typical IMRT radiotherapy treatment plan with seven to nine beam gantry angles, the major machine parameters to be optimized are the MLC leaf positions and movements. For modeling purposes, the patient's body is discretized into small 3D elements called *voxels*, and each beam is discretized into small 2D elements (in the plane perpendicular to the beam direction) called *beamlets* (Figure A.1). A radiation intensity is associated with each beamlet, thus creating a beamlet intensity map from each gantry angle. The MLC leaf positions/movements are optimized in two phases. The intensities of the beamlets are optimized in Phase 1. In Phase 2, another optimization problem is run to optimize the MLC leaf positions/movements to best reproduce the Phase 1 optimal beamlet intensities (Figure A.1.2). In principle, any beamlet intensity can be reproduced perfectly by leaf movements; however, due to engineering considerations, delivery may take a long time. If the beamlet intensities are smooth (i.e., the intensity of the neighboring beamlets do not vary significantly), the optimal intensities can be reproduced efficiently and delivered in a reasonable amount of time. Therefore, smoothness of the beamlet intensities is usually forced or encouraged in the optimization process. All the proposed optimization problems in this paper deal with the optimization of the beamlet intensities and therefore belong in Phase 1. Phase 2 is carried out in an FDA-approved treatment planning system, using its algorithm and our optimal beamlet intensities.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

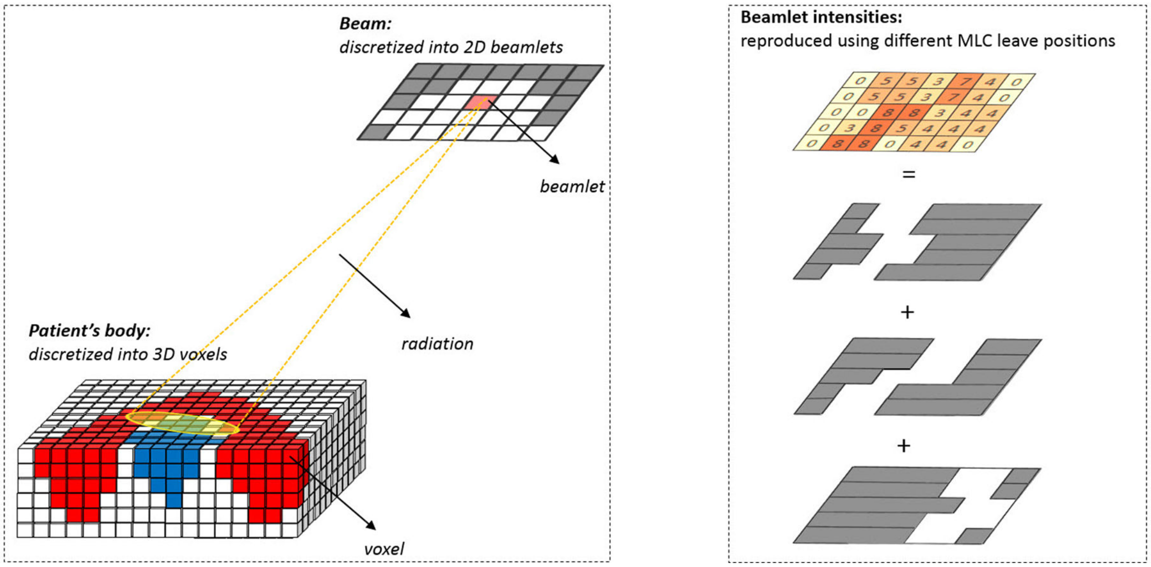


Figure A.1.
In A.1.1, the Problem Is Discretized for Mathematical Modeling; in A.1.2, Beamlet Intensities Are Optimized First and Are Then Realized Using Different Multileaf Collimator (MLC) Leaf Positions
A.1.1: The patient's body and beams are discretized into voxels and beamlets, respectively, and the dose delivered at each voxel from each beamlet is calculated.
A.1.2: MLC leaves move, with the radiation beam on, to build the radiation beams with different shapes and intensities and thus realize the optimal beamlet intensities.

The radiation dose, which is delivered to each voxel $j (j = 1, \ldots, J)$ at the unit intensity of each beamlet $i (i = 1, \ldots, I)$, denoted by a_{ji}, is calculated using a dose calculation algorithm (using the patient's CT scan as a main input). The values a_{ji} define the *dose influence matrix* $A \in R^{J \times I}$, with rows and columns corresponding to the voxels and beamlets, respectively (i.e., beamlets from all candidate beams are concatenated). The number of beamlets is typically on the order of 10,000, whereas the number of voxels is on the order of 100,000. We denote the intensity of the beamlets by $x \in R^I$ and the delivered dose at the voxels by $d \in R^J$. Given the definition of A, there is a linear relationship between the beamlet intensities and the delivered dose: $d = Ax$. The main optimization variables are beamlet intensities $\boldsymbol{x}$; the optimization problem is to search for $\boldsymbol{x}$ with the most desirable delivered dose d. As we mention above, the beamlet intensities of each beam should also be smooth so they can be reproduced efficiently later using MLC leaf movements.

Each voxel belongs to an organ $s \in S$. The influence matrix corresponding to each organ is denoted by A^s ($A^s x$ represents the dose delivered to the voxels of organ s). We assume that a function $f_s(A^s x)$ measures the quality of the delivered dose at organ $\boldsymbol{s}$, and the smaller the value, the better. For example, for a healthy organ, it could be a maximum or mean dose delivered to the organ; for the tumor, it could be the accumulated quadratic deviation of the dose to the prescription dose for all tumor voxels. Assume that a function $h(x)$ measures the quality of the beamlet intensities in terms of smoothness, and the smaller the value, the better. For example, the intensity variation for each beamlet can be calculated as the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

accumulated deviation of that beamlet intensity from all its neighboring beamlets. Then, $h(x)$ can be defined as the accumulated intensity variation for all the beamlets. The general optimization problem then takes the following form:

$$Min\left(f_1\left(A^1x\right), \ldots, f_S\left(A^Sx\right), h(x)\right), \ s.t. \quad x \geq 0 .$$

This is therefore a multicriteria optimization problem with physical nonnegativity constraints on the beamlet intensities. The main challenge is finding a good trade-off between different objectives. The common weighted-sum method does this by minimizing the weighted sum of the objectives; however, it requires manual tuning of the weights in a trial-and-error fashion.

ECHO Mathematical Formulation

Figure A.2 includes the detailed mathematical formulation of the optimization problems. Table A.1 provides a full description of the notations used in the optimization problems. In Step 1, the optimization problem searches for the best tumor irradiation by minimizing the deviation of the dose from the prescription dose while respecting all the maximum and mean dose constraints on tumor and healthy organs. The results obtained in Step 1 are preserved by converting the objective function into a constraint for Step 2, with a slight relaxation parameter to increase the search space in the subsequent optimization step. In Step 2, the healthy organs' doses are minimized subject to the maximum and mean dose constraints of Step 1, plus the constraint added to preserve the results of Step 1. The smoothness of the optimal beamlet intensities is guaranteed by constraining $\boldsymbol{h}(\boldsymbol{x})$ in both optimization problems, and further improvement is also encouraged by adding $\boldsymbol{h}(\boldsymbol{x})$ into the objective function of both problems. Function $\boldsymbol{h}(\boldsymbol{x})$ measures the total variation in beamlet intensities as follows:

$$\mathrm{h}(\mathrm{x}) = \mathrm{w}_1 \sum_{\mathrm{b} \in \mathrm{B}} \left(\sum_{\mathrm{i} \in \mathrm{I_b}} \left(\mathrm{x_i} - \mathrm{x_{R_i}}\right)^2 \right)^2 + \mathrm{w}_2 \sum_{\mathrm{b} \in \mathrm{B}} \left(\sum_{\mathrm{i} \in \mathrm{I_b}} \left(\mathrm{x_i} - \mathrm{x_{L_i}}\right)^2 \right)^2$$

where B is the beam indices, R_i/L_i is the beamlet index of the right/lower neighbor of beamlet i, I_b is the beamlet indices of beam b, and w_1 and w_2 are the smoothing weights. The first term in $\boldsymbol{h}(\boldsymbol{x})$ measures the total variation in the X direction (i.e., leaf movement direction) and the second term accounts for the Y direction. Given that the smoothness in the leaf movement direction is more important, we therefore set $w_1 > w_2$ ($w_1 = 0.6$, $w_2 = 0.4$ in our implementation).

After Step 1 and Step 2 are solved, the optimal beamlet intensities are imported into an FDA-approved treatment planning system for accurate final-dose calculation and to compute the discrepancy between the final dose and the optimization dose (Δ). The correction step (Step C) is then solved to account for discrepancies and correct them. The Lagrange problem (solved in Step C) represents both Steps 1 and 2 for the optimization problem. The Lagrange counterpart is modified slightly ($\boldsymbol{Ax}$ replaced by $\boldsymbol{Ax} + \boldsymbol{\Delta}$) to account for the dose discrepancy.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Constraint (C.1) is added to limit the search space to the vicinity of $\boldsymbol{x^{II}}$ so that Δ is still a valid dose discrepancy.

Step 1 optimization problem:

$$\text{Min } F_I(x) = \sum_{s \in \{\text{Tumor}\}} ||A^s x - p||_2^2 + h(x)$$

$(1.a)\ \max(A^s x) \leq d_s^{max}$: $s \in$ Structures with Max. Dose Constraints

$(1.b)\ \text{mean}(A^s x) \leq d_s^{mean}$: $s \in$ Structures with Mean Dose Constraints

$(1.c)\ h(x) \leq U^h$: Smoothness Constraint

$(1.d)\ x \geq 0$: Fluence Nonnegativity Constraint

Step 2 optimization problem:

$$\text{Min } F_{II}(x) = \sum_{\substack{s \in \text{Healthy} \\ \text{Organs}}} \left|\left|A^s x\right|\right|_{q^s} + h(x), \text{ where } ||d||_q = \left(\frac{\sum_{i \in I} (d_i)^q}{\text{card}(d)}\right)^{\frac{1}{q}}$$

$(2.a)\ \max(A^s x) \leq d_s^{max}$: $s \in$ Structures with Max. Dose Constraints

$(2.b)\ \text{mean}(A^s x) \leq d_s^{mean}$: $s \in$ Structures with Mean Dose Constraints

$(2.c)\ ||(A^s x - p)\,||_2^2 \leq \eta\ ||(A^s x^I - p)\,||_2^2$: $s \in \{\text{Tumor}\}$

$(2.d)\ h(x) \leq U^h$: Smoothness Constraint

$(2.e)\ x \geq 0$: Fluence Nonnegativity Constraint

Step C optimization problem (dose correction step):

$$\text{Min } F_C(x) = F_{II}(x) + \sum_{k \in K} \lambda_k \times g_k(Ax + \Delta)$$

$(C.1)\ \left|\left|x - x^{II}\right|\right|_2^2 \leq \epsilon$

$(C.2)\ x \geq 0$

Figure A.2.

In the Figure, We Provide the Detailed Mathematical Formulations of ECHO's Steps

Table A.1.

The Table Lists the Notations We Use in Our Mathematical Formulation

Symbol	Description
x	Fluence map
p	Prescription dose
A	Influence matrix
s	Structure index
A^s	Influence matrix corresponding to structure s
d_s^{max}	Maximum dose limit for structure s

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Symbol	Description
d_s^{mean}	Mean dose limit for structure s
x^I, x^{II}	Step 1, Step 2 optimal solution
q^s	q parameter for structure s
η	Slip parameter
U^h	Upper bound for total variation of beamlet intensities
g_k	Constraint k of Step 2
λ_j	Lagrange multiplier of Constraint j
Δ	The discrepancy between the dose from Step 2 and final dose calculation
ϵ	Correction step search space constraint
card(d)	Cardinality of d (i.e., the number of elements)

Appendix B: Challenges and OR Solutions (in Research and Development)

In this appendix, we review three challenges and their successfully prototyped OR solutions. We have not yet implemented these solutions in clinical practice.

Challenge 1: Continuous Delivery of Radiation

Challenge:

IMRT delivers radiation from a few beam angles. In a more recent radiation delivery technique, volumetric modulated arc therapy (VMAT), the gantry rotates and delivers radiation continuously. Because the beam is always on in this technique, it provides a shorter delivery time. VMAT is approximately twice as fast as standard IMRT and is a method of choice for treating tumors in some disease sites. However, VMAT represents a much larger and more challenging optimization problem than IMRT. On one hand, VMAT optimization involves many more beams than IMRT because of the continuous delivery of radiation. On the other hand, the continuous gantry rotation does not allow the beam intensity modulation at each beam; therefore, unlike IMRT, we cannot optimize the beamlet intensities first and convert them to apertures later. VMAT requires the optimization of leaf positions directly, which has a nonconvex relationship with the delivered radiation dose to the patient body.

Solution:

Our research meets this challenge using SCP. The idea is to solve this nonconvex problem as a sequence of simplified convex approximations. The main challenge is how to create these approximations by taking advantage of the special structure of the problem. We generate convex approximations in our application by restricting the collimator leaf movements at each iteration (Dursun et al. 2021).

SCP is a heuristic algorithm, as we illustrate in Figure B.1. We validated its performance by comparing the result against the ground truth obtained using an MINLP on a simple and aggressively down-sampled problem for which we can computationally afford to solve the resultant MINLP problem. Figure B.2 shows that SCP generates results that are similar to the ground truth but in a fraction of time (~2 minutes versus ~25 minutes).

Author Manuscript

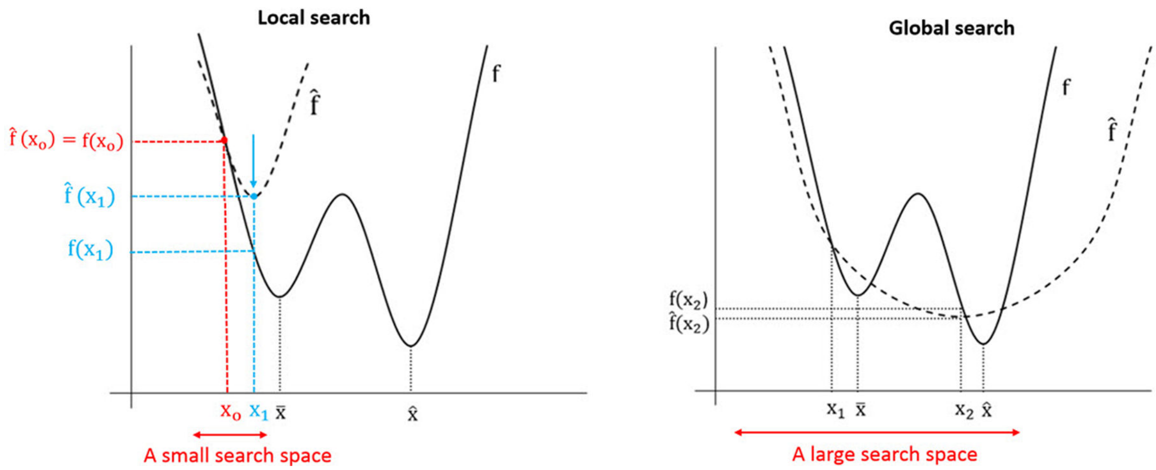


Figure B.1.

The Schematic Illustrates the Concept Underlying Sequential Convex Programming

Notes. In the figure on the left, a one-dimensional nonconvex function f (solid line) is approximated locally at x_0 by a convex function $\hat{f}$ (dashed line). The optimal solution of $\hat{f}$ is found (x_1) and used as the next solution. In the figure on the right, a nonconvex function f is approximated globally at x_1 by $\hat{f}$ whose optimal solution x_2 is used as the next solution.

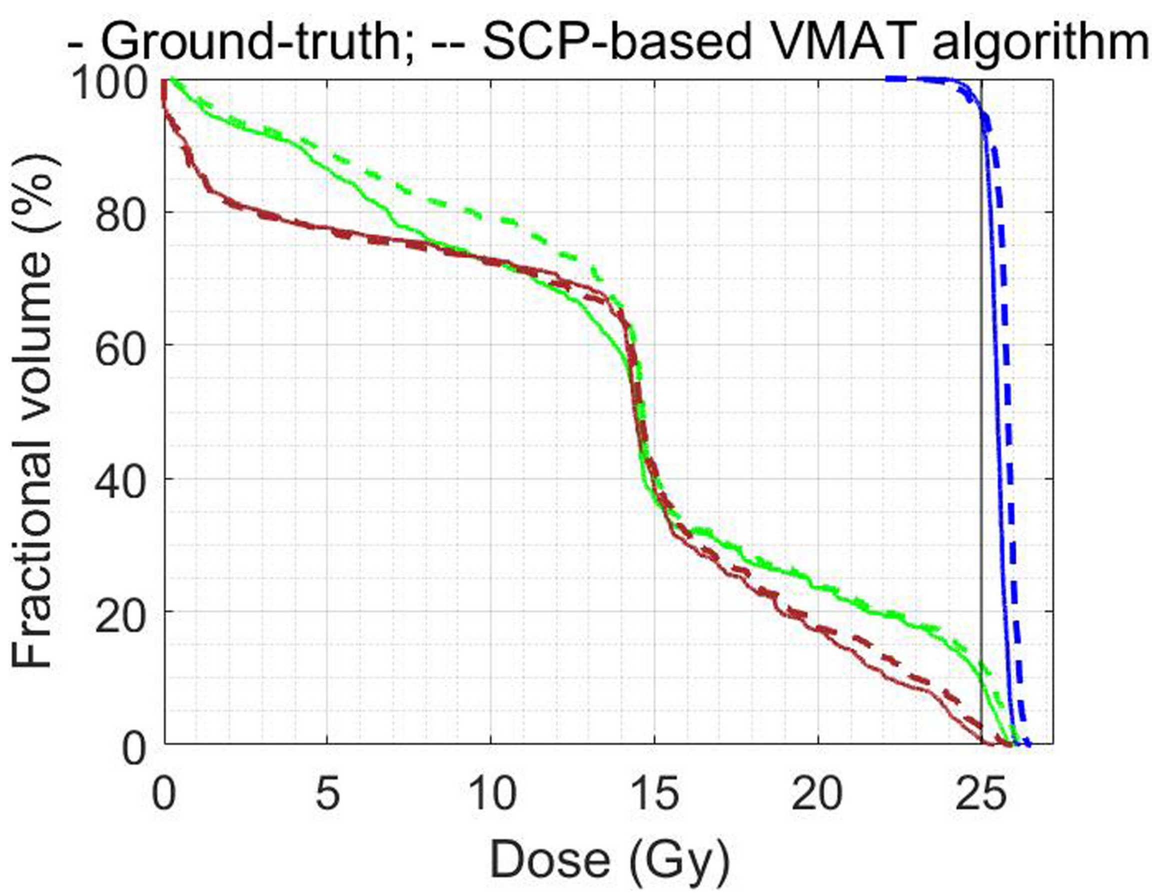


Figure B.2.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The Effectiveness of the Sequential Convex Programming Technique to Handle Nonconvex VMAT Optimization Is Illustrated by Comparing Its Results Against the Ground Truth Obtained by Solving the Corresponding MINLP Problem

Novelty:

To the best of our knowledge, this is the first VMAT optimization technique equipped with a global search strategy to promote the global convergence and provide a near-global optimal solution. The gradient-based technique, which is the basis of many current commercial implementations of VMAT (Unkelbach et al. 2015) suffers from the local optimality and the quality of the plan heavily relies on providing a good initial point, whereas our algorithm is fairly insensitive to the initialization, as shown in our experiments. This is also the first use of SCP for radiotherapy optimization and the first time the quality of a heuristic algorithm of VMAT is verified by comparing it against the ground truth solution.

Challenge 2: Beam Angle Selection for Proton Therapy

Challenge:

The selection of the beam angles from which the radiation is delivered to the patient is another interesting and challenging problem. It is especially important in proton therapy, where high-energy protons are used instead of photons as a source of radiation. In this modality, the treatment is usually delivered from only two to three beam directions. Choosing two to three beams from the large number of possibilities is a challenging combinatorial optimization problem.

Solution:

Bayesian optimization, which we use to address this challenge, is a global optimization technique suitable for smooth complex functions without an explicit expression, as is the case for beam angle selection problems. For this problem, the quality of a set of beams is only realized after solving the corresponding optimization problem; therefore, the optimization function is unknown. The key concept in Bayesian optimization is to leverage the points at which the function has already been evaluated to create an estimate of the unknown function to guide the optimization. For our application, we run ECHO for some initial beam angle candidates and rate the resultant treatment plan for each beam configuration using a clinically relevant treatment-score function. Bayesian optimization iteratively predicts the treatment-score candidates that have not yet been evaluated to find the best candidate to be optimized next with ECHO. Our experiments on five head-and-neck patients show that by integrating ECHO with Bayesian optimization, we can find the optimal beam selection by evaluating at most only 4% of all potential beam configurations (Taasti et al. 2020b).

Novelty:

For the conventional weighted-sum method, different optimization schemes have been suggested for beam angle selection, including the exhaustive search method (Meedt et al. 2003), and mixed-integer programming (Lee et al. 2003b, Lim et al. 2008). To the best of our knowledge, this is the first automated treatment planning system with beam angle selection and also the first application of Bayesian optimization in radiotherapy treatment planning.

Challenge 3: Uncertainty Management for Proton Therapy

Challenge:

The outcome of radiotherapy treatment can be compromised by unavoidable uncertainties, such as patient setup errors, during the treatment. The affects are even more pronounced for proton therapy compared with the use of photons due to the sharp dose fall-off of protons.

Solution:

We equip ECHO with a robust optimization platform by simulating various errors and then optimizing the treatment plan with the objective of achieving a good treatment plan even in the presence of error scenarios. We combine the objective functions of the individual scenarios using the p-norm function. The p-norm with a parameter $p = 1$ or $p = \infty$ results in the stochastic or the worst-case approach, respectively; an intermediate robustness level is obtained by employing p-values between 1 and ∞.

In the example we show in Figure B.3, the uncertainty spread is much smaller for the robust plans in pink, compared with the nonrobust plans in blue.

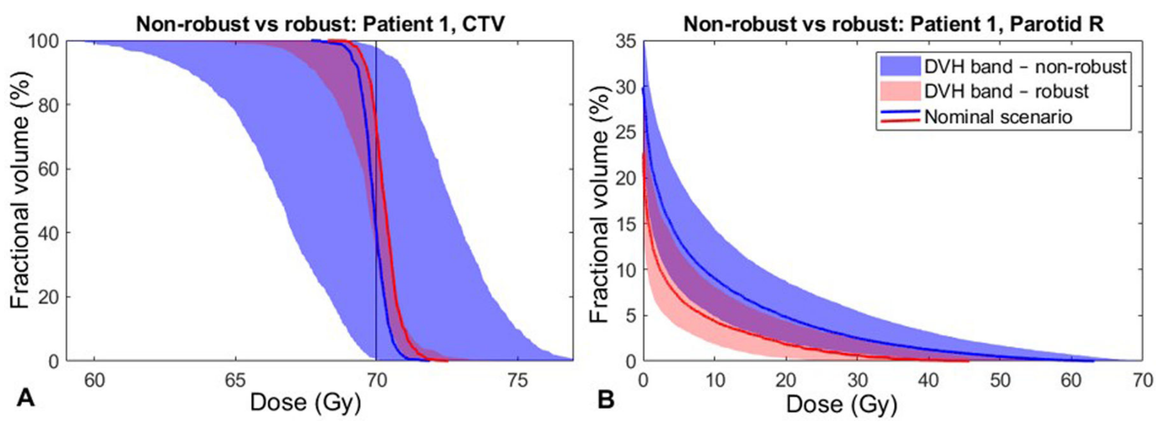


Figure B.3.
The Figure Illustrates Bands of Possible Delivered-Dose Volume Curves for the Tumor (Left Figure) and Parotid Gland (Right Figure) for Robust (Red) and Nonrobust (Blue) Plans
Note. Th**e** P-norm function with p = 2 is used in robust optimization.

Novelty:

Robust optimization has been introduced in the past for the conventional trial-and-error weighted-sum method (Unkelbach et al. 2018), however, this is the first work that addresses automated robust planning. Moreover, this work is novel in terms of using p-norm to adjust the level of robustness and fill the gap between the two common extreme robust optimization approaches (i.e., stochastic and worst-case methods).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

References

Appenzoller LM, Michalski JM, Thorstad WL, Mutic S, Moore KL (2012) Predicting dose-volume histograms for organs-at-risk in IMRT planning. Medical Phys. 39(12):7446–7461. 10.1118/1.4761864.

Bortfeld T (2006) IMRT: A review and preview. Phys. Medicine Biol 51(13):R363–R379. 10.1088/0031-9155/51/13/R21.

Brahme A, Roos JE, Lax I (1982) Solution of an integral equation encountered in rotation therapy. Phys. Medicine Biol 27(10):1221–1229. 10.1088/0031-9155/27/10/002.

Breedveld S, Storchi PRM, Heijmen BJM (2009) The equivalence of multi-criteria methods for radiotherapy plan optimization. Phys. Medicine Biol 54(23):7199–7209.

Clark VH, Chen Y, Wilkens J, Alaly JR, Zakaryan K, Deasy JO (2008) IMRT treatment planning for prostate cancer using prioritized prescription optimization and mean-tail-dose functions. Linear Algebra Appl. 428(5–6):1345–1364. 10.1016/j.laa.2007.07.026. [PubMed: 18974791]

Craft D, Bortfeld T (2008) How many plans are needed in an IMRT multi-objective plan database? Phys. Medicine Biol 53(11):2785–2796. 10.1088/0031-9155/53/11/002.

Deasy JO (2002) The IMRT optimization problem statement. Proc. NCI-NSF Sponsored Workshop Oper. Res. Appl. Radiation Therapy (ORART) (National Cancer Institute at the National Institutes of Health, Washington, DC–National Science Foundation, Alexandria, VA).

Deasy JO (2000) Prioritized treatment planning for radiotherapy optimization. Proc. World Congress Medical Phys. Biomedical Engr CD-ROM (Chicago).

Dursun P, Zarepisheh M, Jhanwar G, Deasy JO (2021) Solving the volumetric modulated arc therapy (VMAT) problem using a sequential convex programming method. Phys. Med. Biol 66(8):x–y. 10.1088/1361-6560/abee58.

Ehrgott M, Güler Ç, Hamacher H, Shao L (2010) Mathematical optimization in intensity modulated radiation therapy. Ann. Oper. Res 175(November):309–365. 10.1007/s10479-009-0659-4.

Ge Y, Wu QJ (2019) Knowledge-based planning for intensity-modulated radiation therapy: A review of data-driven approaches. Medical Phys. 46(6):2760–2775. 10.1002/mp.13526.

Hong L, Zhou Y, Yang J, Mechalakos JG, Hunt MA, Mageras GS, Yang J, Yamada J, Deasy JO, Zarepisheh M (2019) Clinical experience of automated SBRT paraspinal and other metastatic tumor planning with constrained hierarchical optimization. Advances Radiation Oncology 5(5):1042–1050. 10.1016/j.adro.2019.11.005.

Langer M, Brown R, Urie M, Leong J, Stracher M, Shapiro J (1990) Large scale optimization of beam weights under dose-volume restrictions. Internat. J. Radiation Oncology Biol. Phys 18(4):887–893.

Lee EK, Deasy J, Langer M, Rardin R, Deye JA, Mahoney FJ (2003a) NCI/NSF workshop on operations research applied to radiation therapy Washington, DC, February 7–9, 2002. Ann. Oper. Res 119(March):143–146. 10.1023/A:1022982407025.

Lee EK, Fox T, Crocker I (2003b) Integer programming applied to intensity-modulated radiation therapy treatment planning. Ann. Oper. Res 119(March):165–181. 10.1023/A:1022938707934.

Leibel SA, Fuks Z, Zelefsky MJ, Wolden SL, Rosenzweig KE, Alektiar KM, Hunt MA, et al. (2002) Intensity-modulated radiotherapy. Cancer J. 8(2):164–176. [PubMed: 12004802]

Lim GJ, Choi J, Mohan R (2008) Iterative solution methods for beam angle and fluence map optimization in intensity modulated radiation therapy planning. OR Spectrum 30(July):289–309. 10.1007/s00291-007-0096-1.

Meedt G, Alber M, Nüsslin F (2003) Non-coplanar beam direction optimization for intensity-modulated radiotherapy. Phys. Medicine Biol 48(September):2999–3019. 10.1088/0031-9155/48/18/304.

Monz M, Küfer KH, Bortfeld TR, Thieke C (2008) Pareto navigation—algorithmic foundation of interactive multi-criteria IMRT planning. Phys. Medicine Biol 53(September):985–998. 10.1088/0031-9155/53/4/011.

Mukherjee S, Hong L, Deasy JO, Zarepisheh M (2020) Integrating soft and hard dose-volume constraints into hierarchical constrained IMRT optimization. Medical Phys. 47(2):414–421. 10.1002/mp.13908.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Romeijn HE, Dempsey JF (2008) Intensity modulated radiation therapy treatment plan optimization. TOP 16(November):215–243. 10.1007/s11750-008-0064-1.

Shepard DM, Ferris MC, Olivera GH, Mackie TR (1999) Optimizing the delivery of radiation therapy to cancer patients. SIAM Rev. 41(4):721–744. 10.1137/S0036144598342032.

Siebers JV, Lauterbach M, Tong S, Wu Q, Mohan R (2002) Reducing dose calculation time for accurate iterative IMRT planning. Medical Phys. 29(2):231–237. 10.1118/1.1446112.

Taasti VT, Hong L, Deasy JO, Zarepisheh M (2020a) Automated proton treatment planning with robust optimization using constrained hierarchical optimization. Medical Phys. 47(7): 2779–2790. 10.1002/mp.14148.

Taasti VT, Hong L, Shim JSA, Deasy JO, Zarepisheh M (2020b) Automating proton treatment planning with beam angle selection using Bayesian optimization. Medical Phys 47(8):3286–3296. 10.1002/mp.14215.

Unkelbach J, Alber M, Bangert M, Bokrantz R, Chan TCY, Deasy JO, Fredriksson A, et al. (2018) Robust radiotherapy planning. Phys. Medicine Biol 63(22):22TR02. 10.1088/1361-6560/aae659.

Unkelbach J, Bortfeld T, Craft D, Alber M, Bangert M, Bokrantz R, Chen D, et al. (2015) Optimization approaches to volumetric modulated arc therapy planning. Medical Phys. 42(3):1367–1377. 10.1118/1.4908224.

Varian (2021) Eclipse. Accessed July 19, 2021, https://www.varian.com/products/radiotherapy/treatment-planning/eclipse.

Webb S (1989) Optimisation of conformal radiotherapy dose distribution by simulated annealing. Phys. Medicine Biol 34(10):1349–1370. 10.1088/0031-9155/34/10/002.

Wilkens JJ, Alaly JR, Zakarian K, Thorstad WL, Deasy JO (2007) IMRT treatment planning based on prioritizing prescription goals. Phys. Medicine Biol 52(6):1675–1692. 10.1088/0031-9155/52/6/009.

Zarepisheh M, Deasy JO, Hong L, Mageras GS, Hunt MA, Mechalakos JG (2019a) Methods and systems for automatic radiotherapy treatment planning. World Intellectual Property Organization patent application WO2019173625A1, published December 9, 2019.

Zarepisheh M, Hong L, Zhou Y, Oh JH, Mechalakos JG, Hunt MA, Mageras GS, Deasy JO (2019b) Automated intensity modulated treatment planning: The expedited constrained hierarchical optimization (ECHO) system. 46(7):2944–2954. 10.1002/mp.13572.

Author Manuscript

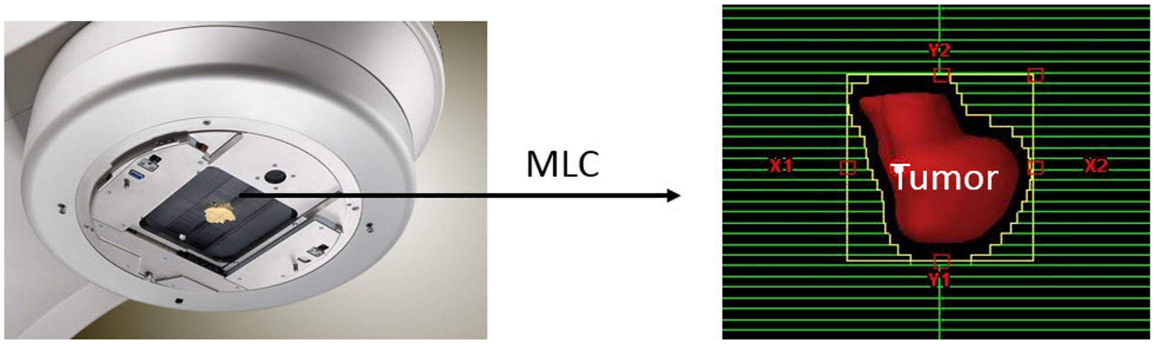


Figure 1.

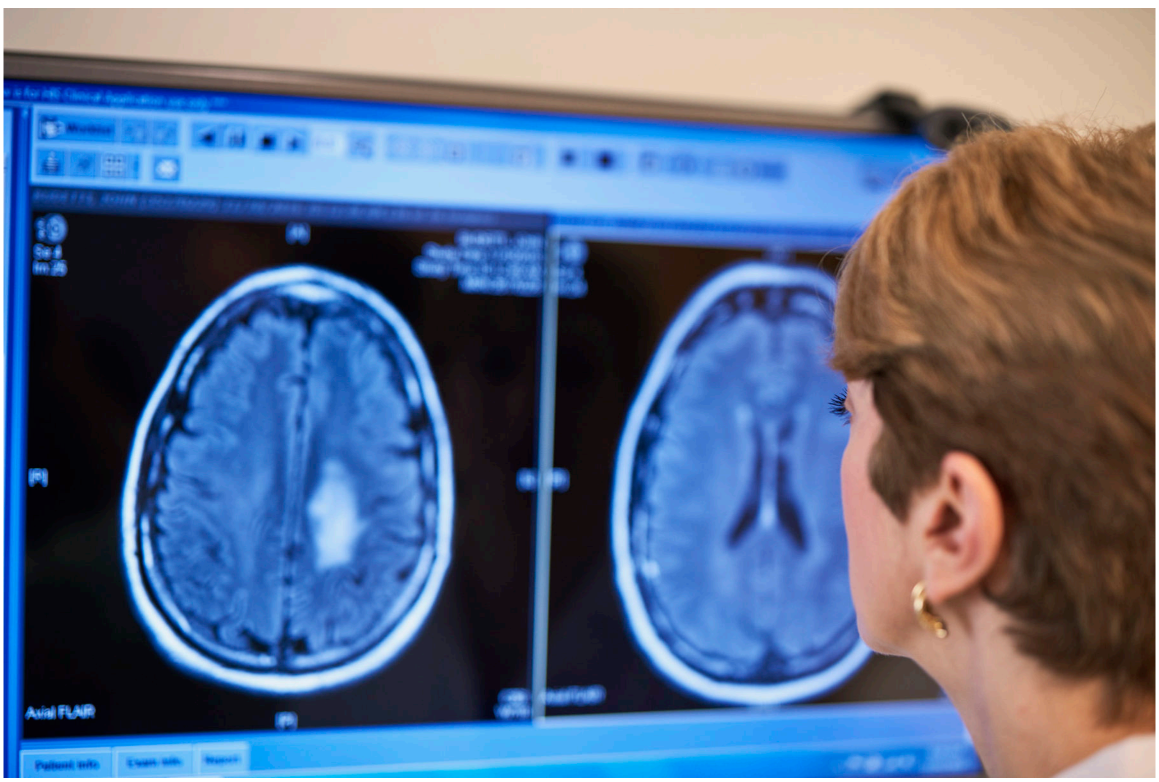
The Figures Illustrate the Process by Which Multileaf Collimators Work

1A: The Multileaf collimator (MLC) is mounted on the linear accelerator gantry's head, facing the patient.

1B: A "beams eye view" provides a closer look at the MLC. A typical MLC has 60 tungsten leaf pairs with each leaf width between 0.25cm and 1.0 cm. The total MLC size projected to the patient ranges from 20cm to 40cm.

Notes. The radiation delivery machine shapes the beams by passing them through the MLC (Figures 1A and 1B). The leaves of an MLC move independently and typically form many different open patterns that sum up to a modulated intensity profile for that gantry angle (beam).

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript



Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Figure 2.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The Pictures, Which Were Taken at MSK, Illustrate the Four Main Steps of Radiation Therapy

2A. Imaging: the patient's digital image is acquired using an imaging device (e.g,. CT, MRI, PET).

2B. Contouring: A physician defines and draws the borders of the tumor and healthy organs on the digital image.

2C. Treatment Planning: An expert planner (e.g., dosimetrist or physicist) uses software to develop the plan and customize the machine parameters. A physician (radiation oncologist) reviews the final plans.

2D. Delivery: The delivery machine rotates around the patient and delivers radiation from different directions.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

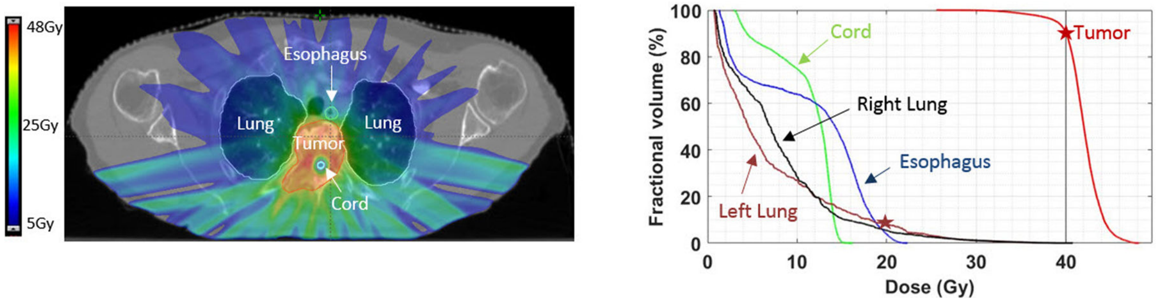


Figure 3.

The Graphics Illustrate Dose Distribution (3A) and a Dose Volume Histogram (DVH) (3B), Which Are the Two Main Plan Evaluation Tools

3A. Dose distribution is a three-dimensional color-coded map showing how much radiation is delivered to different organs. The high-dose radiation is focused on the tumor.

3B. DVH is a two-dimensional plot, with a curve for each organ, specifying the fraction of the volume of the organ receiving at least a specified amount of dose.

Note. The prescription dose of 40 gray or Gy (a unit of medical radiation expressing in joules per kilogram the absorbed energy per unit of mass tissue) is desired at the tumor for this illustrative prostate example.

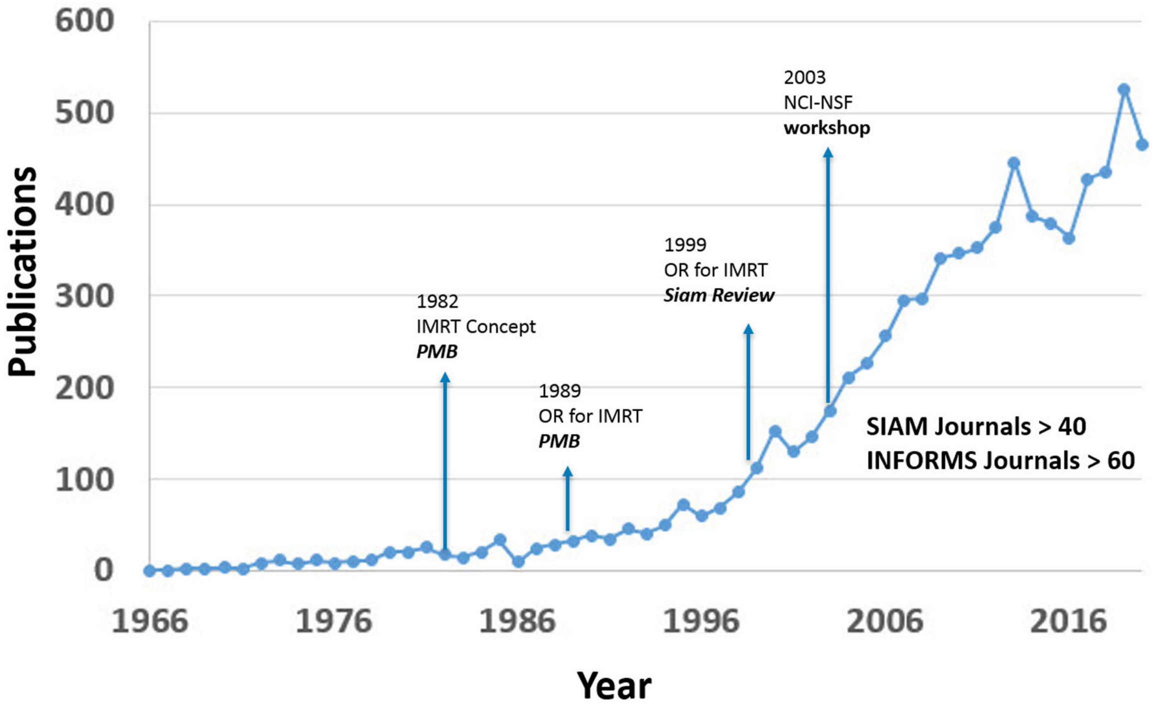


Figure 4.

The Graph Shows the Annual Number of OR Publications on Radiotherapy and Highlights Important Publications and Events

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

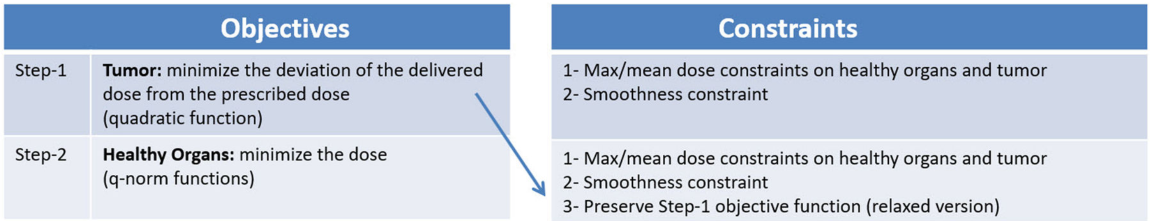


Figure 5.

The Graphic Shows a High-level Overview of the ECHO Framework

Note. The arrow indicates how the Step 1 optimization results are carried over to Step 2 as slightly relaxed constraints.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

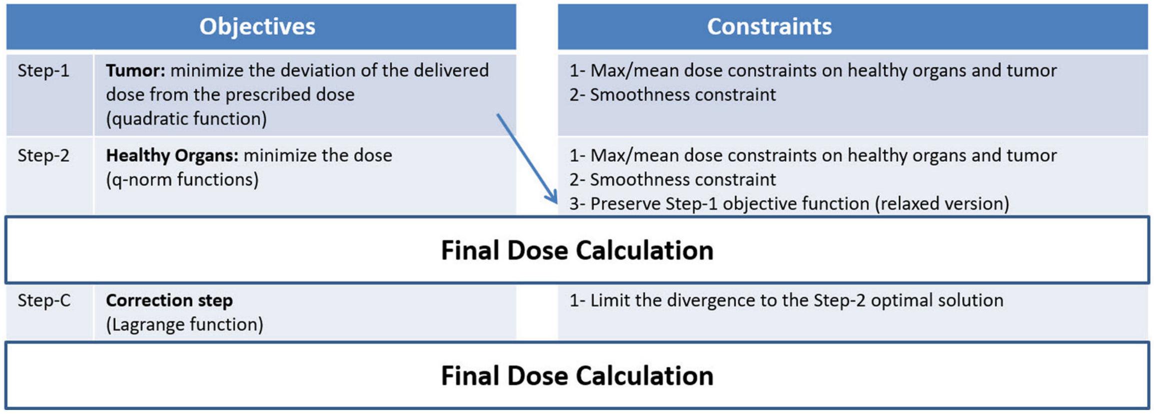


Figure 6.

The Correction Step (Step C) Adjusts For Discrepancies Between the Optimization Results and the Final-dose Calculation, Which Is Performed Within an FDA-approved Treatment Planning System

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Objectives		Constraints
Step-0	**Tumor:** minimize the deviation of the delivered dose from the prescribed dose (quadratic function)	1- Max/mean dose constraints on healthy organs and tumor 2- Smoothness constraint 3- Convex approximation constraints
Step-1	**Tumor:** minimize the deviation of the delivered dose from the prescribed dose (quadratic function)	1- Max/mean dose constraints on healthy organs and tumor 2- Smoothness constraint 3- Max dose constraint on low-dose voxels
Step-2	**Healthy Organs:** minimize the dose (q-norm functions)	1- Max/mean dose constraints on healthy organs and tumor 2- Smoothness constraint 3- Preserve Step-1 objective function (relaxed version) 4- Max dose constraint on low-dose voxels

Figure 7.

We Use a Heuristic Technique to Handle Nonconvex DVH Constraints by Adding Step 0 to the Problem Using a Convex Approximation of the Nonconvex Constraints and by Adding Maximum Dose Constraints on Lose-dose Voxels in Step 2 and Step 3

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

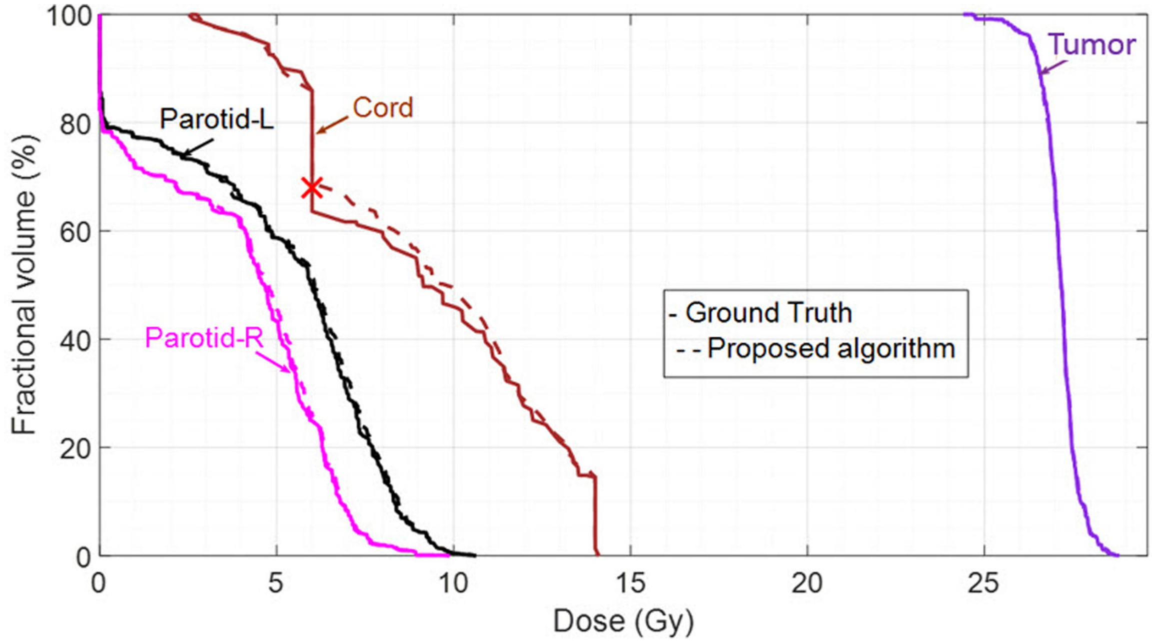


Figure 8.

The Effectiveness of the Heuristic Technique Used to Handle Nonconvex DVH Constraints Is Illustrated by Comparing Its Results to the Ground Truth Obtained by Solving the MINLP Problem

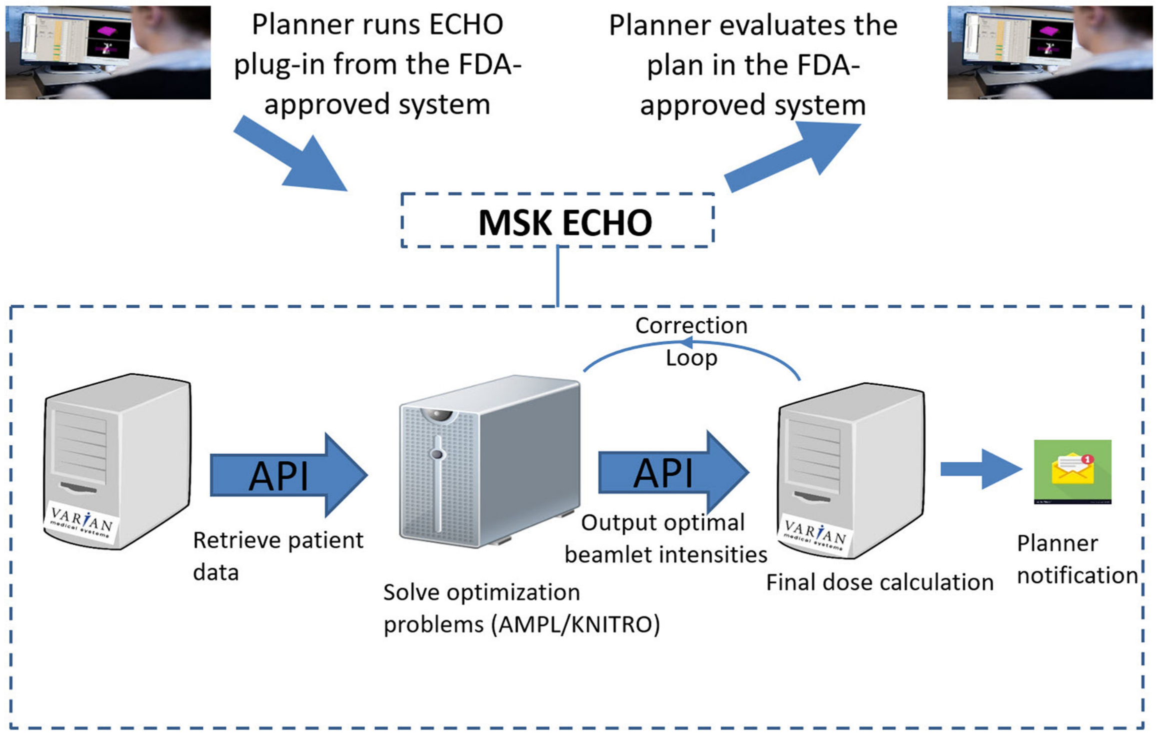


Figure 9.

The Graphic Illustrates the Integration of the ECHO System with an FDA-approved Treatment Planning System Using API Scripting

Notes. A planner runs ECHO from an FDA-approved system and all the patient data are transferred to MSK servers for optimization. The optimal beam intensities are imported back to the clinical system for final dose calculations and plan evaluation.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

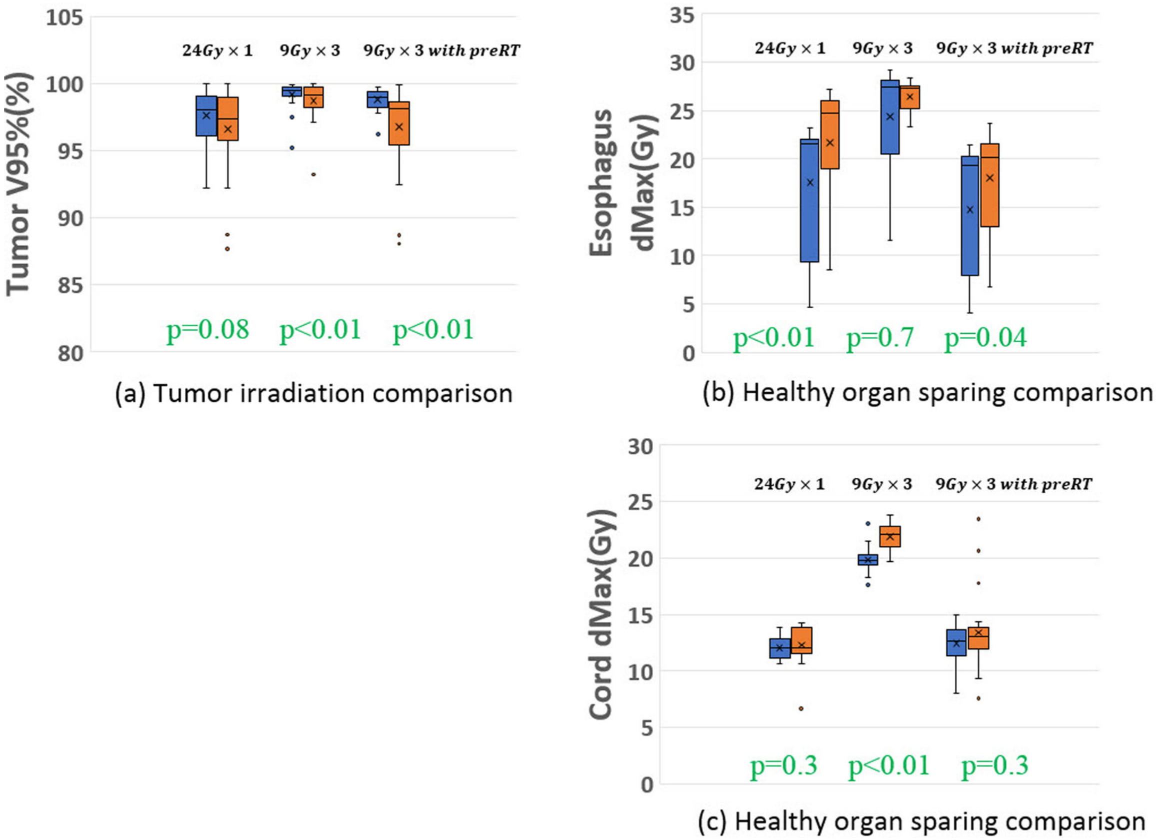


Figure 10.

The Graphic Shows a Comparison of Automated ECHO Plans (Blue Bars on the Left) and Manually Created Clinical Plans (Orange Bars on the Right) for 75 Paraspinal Plans with Three Prescription Schemes (25 Plans Per Scheme)

Notes. As the left plot shows, ECHO delivers more radiation to the tumor (a higher value is better in (a)); as the right plot shows, ECHO delivers less radiation to the healthy organs (a lower value is better in (b) and (c)). For statistical tests, we used a Wilcoxon signed-rank test and considered $p < 0.05$ as statistically significant. For example, for a prescribed treatment dose of $24Gy \times 1$ in (a), ECHO (blue bar) delivers at least 95% of the prescription dose to a larger volume of the tumor compared with the manual plan (orange bar) with statistical confidence intervals of p=0.08.

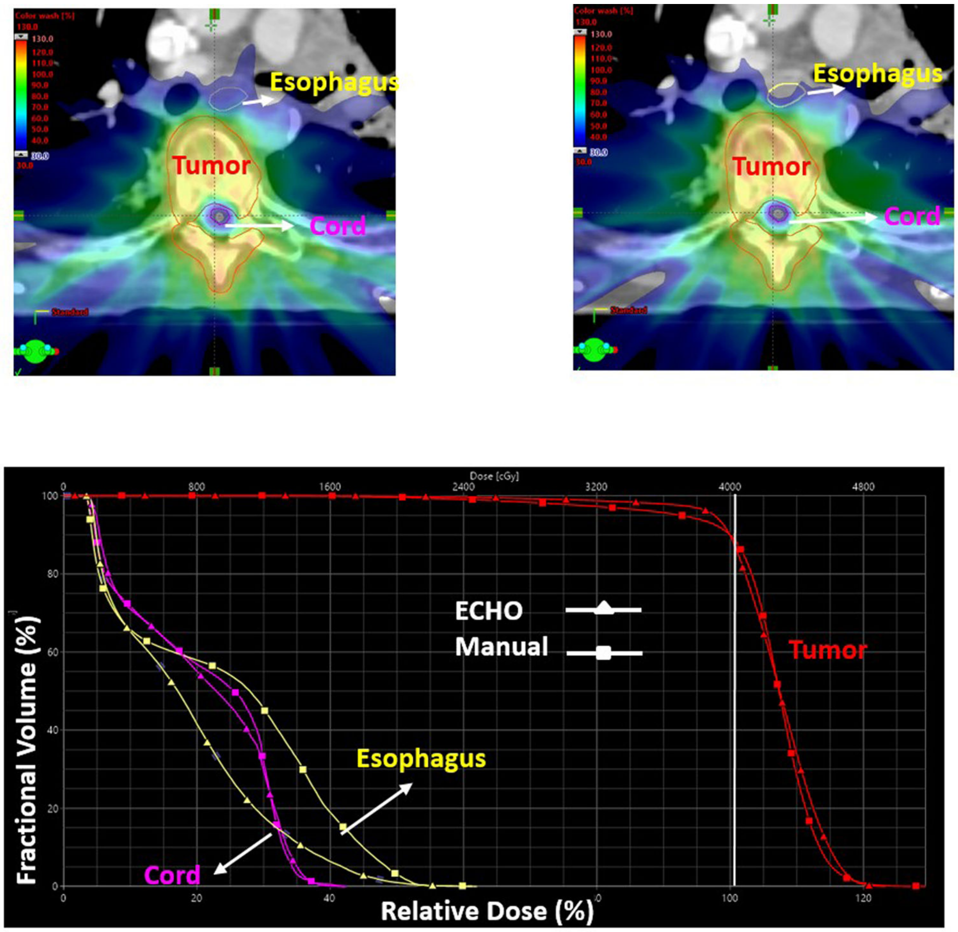


Figure 11.

The DVH and Dose Distribution Reveal that an Automated ECHO Plan Preserves More of the Esophagus than a Clinical, Manually Created Plan

A: Dose distribution of the clinical manual plan

B: Dose distribution of the ECHO plan

Note. The white horizontal line on the DVH (Relative Dose %) represents the prescription.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

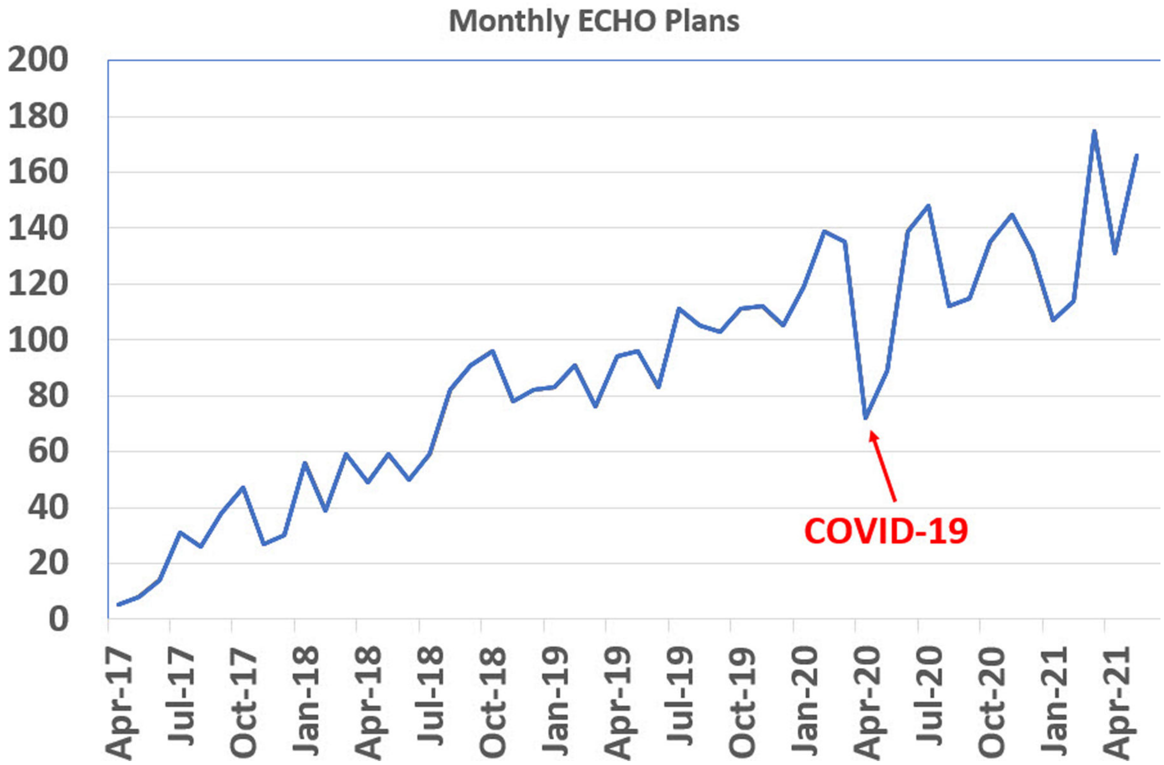


Figure 12.

The Graphic Shows the Monthly Number of ECHO Plans Used at MSK Since ECHO's Inception

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

The Table Provides a Summary of the Challenges, OR Solutions, and Corresponding Publications

Implementation Status	Challenge	OR Solution	Publication
Clinically Implemented	Conflict between tumor irradiation and healthy tissue sparing	Hierarchical constrained nonlinear optimization	(Zarepisheh et al. 2019b)
	1- Solving large-scale constrained optimization problems within restrictive clinical time frames 2 - Modeling inaccuracies	Solving a problem with a truncated influence matrix first and accounting for inaccuracies later using Lagrangian methods	
	3 - Nonconvex clinical criteria	Mixed-integer nonlinear programming combined with an effective heuristic	(Mukherjee et al. 2020)
Research/Development	Continuous delivery of radiation (VMAT)	Sequential convex programming (SCP)	(Dursun et al. 2021)
	Beam angle selection for proton therapy	Bayesian optimization	(Taasti et al. 2020b)
	Management of treatment uncertainties for proton therapy	Robust optimization with p-norm function to control the level of robustness	(Taasti et al. 2020a)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

We Summarize the Benefits Provided by ECHO Compared With the Clinical Practice Before ECHO's Implementation

Benefit category	Comparison of ECHO with the clinical practice	Benefits
Cancer-care quality	Faster	Shorter time to treatment (shortening the average full-treatment planning time, which includes plan generation and QA processes for nonexpedited cases, by one day, from five days to four days).
		Avoidance of unnecessary surgeries
		Enabling of same-day treatment for patients in urgent need
	Higher quality	Expectation of less radiation-induced side effects (e.g., 18% less maximum dose to esophagus on average for spine patients treated with 24Gy × 1)
		Expectation of better disease control (e.g., 2.5% increase in the fraction of tumor volume receiving at least prescription dose (2.5% increase in V(100%))
	Less variability	High-quality plan regardless of planner's experience (e.g., tumor V(100%) ranged between 91% and 100% with ECHO versus a range of 79% to 100% with manual plans)
Financial/Resource	Faster	Shorter hospital stays for patients
	Less expensive	Increased productivity by at least 15% per planner when using ECHO