Monte Carlo-based inverse treatment planning

This article has been downloaded from IOPscience. Please scroll down to see the full text article.


(http://iopscience.iop.org/0031-9155/44/8/303)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 129.78.32.23
The article was downloaded on 10/05/2012 at 16:16

Please note that terms and conditions apply.
Monte Carlo-based inverse treatment planning

Robert Jeraj† and Paul Keall‡§

† Reactor Physics Division, Jožef Stefan Institute, Ljubljana, Slovenia
‡ Department of Radiation Oncology, Medical College of Virginia, Richmond, VA, USA
§ Cancer Therapy Centre, Liverpool Hospital, NSW, Australia

E-mail: robert.jeraj@ijs.si

Received 22 December 1998, in final form 15 March 1999

Abstract. A Monte Carlo-based inverse treatment planning system (MCI) has been developed which combines arguably the most accurate dose calculation method (Monte Carlo particle transport) with a ‘guaranteed’ optimization method (simulated annealing).

A distribution of photons is specified in the tumour volume; they are transported using an adjoint calculation method to outside the patient surface to build up an intensity distribution. This intensity distribution is used as the initial input into an optimization algorithm. The dose distribution from each beam element from a number of fields is pre-calculated using Monte Carlo transport. Simulated annealing optimization is then used to find the weighting of each beam element, to yield the optimal dose distribution for the given criteria and constraints.

MCI plans have been generated in various theoretical phantoms and patient geometries. These plans show conformation of the dose to the target volume and avoidance of critical structures. To verify the code, an experiment was performed on an anthropomorphic phantom.

(Some figures in this article appear in colour in the electronic version; see www.iop.org)

1. Introduction

In conventional treatment planning the number of parameters to be optimized (i.e. position of beams, number and types of wedges) is small and can be performed with a trial and error procedure by an experienced treatment planner. However, with the increase in the degrees of freedom offered by intensity modulated radiation therapy, inverse treatment planning becomes almost a necessity.

There are three main ingredients of every inverse treatment planning algorithm: the dose calculation method, the optimization algorithm and the objective function. The dose calculation should be accurate enough to ensure that the ‘prediction’ of the dose distribution indeed matches the dose delivered to the patient. Most of the dose calculation methods used so far in the inverse treatment planning are based on a pencil-beam convolution or superposition approach (Mackie et al 1993, Bortfeld et al 1993, Holmes et al 1998, Saxner and Ahnesjö 1998). The main drawback of all these methods is how the inhomogeneities in a patient are treated. This is usually done by some scaling methods, which at best account only for particle transport on a macroscopic scale (Mohan 1997, Nahum 1997). Transport on a macroscopic scale leads to errors at interfaces of different densities in a patient body (Woo and Cunningham 1990, Yu et al 1995, Keall and Hoban 1995, 1996). It has been argued that use of Monte Carlo dose calculations may lead into improved clinical outcome of radiotherapy treatment (Mohan 1997, Nahum 1997). However, inverse treatment planning will benefit even more from the use
of Monte Carlo transport than conventional planning. In inverse treatment planning, not only is the accuracy of the dose calculation improved, but also the optimized answer is consequently different. Thus, if an inaccurate dose calculation is used, one might converge to an ‘optimized’ but not optimal solution in the sense that the optimization was guided by a misleading dose calculation. The main disadvantage of Monte Carlo is the extensive computation time required. With the increased speed of computers, attempts have been made to implement Monte Carlo in clinical planning systems (Rogers et al 1990, Hartmann-Siantar et al 1997, Mubata et al 1998, Libby et al 1998, Siebers et al 1998). Calculation speed is less critical for computer optimized planning than manual planning, as no constant human interaction is required. Hence a calculation time even of the order of a day is acceptable, provided that the resultant plan is satisfactory.

The second basic ingredient of the inverse treatment planning is mathematical optimization, which should give a dose distribution that satisfies the constraints given for a specific case (i.e. uniformity of the dose, conformity to the tumour, sparing of the critical tissues). There have been many algorithms applied to the optimization problem in radiotherapy (for a good review see Ebert (1997a, b, c)). These algorithms can be separated into two groups, deterministic and stochastic. Deterministic methods use well-defined (calculated) search directions to determine the next iteration step. The most common method is the gradient-based method (Brahme 1988, Bortfeld et al 1990, Holmes et al 1991) which is known for its speed; however, it may also get trapped in a local minimum (Press et al 1992). This problem is usually not too severe, especially when simple objective functions are used (with no local minima), when a good first approximation is used or when local minima have similar values to the global minimum of the objective function (Bortfeld 1997). The stochastic methods, on the other hand, choose the search direction randomly. The most extensively used stochastic method in radiotherapy optimization has been the simulated annealing method (Webb 1989, 1995, Morrill et al 1991, Sloboda 1992, Ebert and Hoban 1997), though other methods like genetic algorithms have also been applied (Langer et al 1996, Yu 1997). Simulated annealing stochastically searches the space of feasible solutions and with a good cooling-down algorithm ensures convergence to the global minimum. Though simulated annealing can be slow, there are several techniques to speed it up (Szu and Hartley 1987, Mageras and Mohan 1993, Webb 1995, Oldham et al 1995, Oldham and Webb 1995, Rosen et al 1995). Though certain types of objective functions, i.e. least square, may have only one minimum (Bortfeld 1997), it has been shown that for several different types of objective, the introduction of complexity in dose delivery (i.e. multiple beam angles) forces the introduction of local minima (Niemierko 1996, Deasy 1996). Another advantage of simulated annealing is the flexibility in incorporating different objective functions into the algorithm.

Another important constituent of inverse treatment planning, as mentioned above, is the objective function used in the optimization algorithm, either being an explicit function or implicitly taken into account during optimization iterations or just forming a set of constraints that have to be met as closely as possible. The objective function should be flexible such that it allows different requirements to be posed by users, because no agreement has been achieved about whether just physical optimization is enough or use of biological parameters would give better survival outcomes and which form of the objective function is best. The study of different objective functions is not addressed in this article; however, the effect of the function, and of the values of the parameters within the function, on the final plan cannot be understated. For an excellent discussion on the importance of the objective functions and especially the difference between physical and biological optimization readers are referred to Wang et al (1995), Bortfeld et al (1996), Mohan and Wang (1996) and Goitein and Niemierko (1996).
Monte Carlo-based inverse treatment planning

Though Monte Carlo has been applied to radiation therapy, to date no literature has investigated Monte Carlo particle transport for inverse treatment planning. However, without having a better dose calculation incorporated into an inverse treatment planning algorithm it is impossible to say how far from the optimized result such a solution is. As already mentioned, there are two errors which arise through the use of an approximate dose calculation—one is because of the approximate dose calculation itself, which can be significant or not (probably there should not be much difference for example in the abdominal region, but what in the head and neck region?) and the second error which arises because of the convergence to a different ‘optimal solution’. How severe these errors are, one can only determine by using a more accurate dose calculation method. The aim of this work is to develop an inverse treatment planning system that combines the most accurate dose calculation algorithm with a robust optimization algorithm. Such an algorithm can also be used to test other faster but potentially less accurate inverse treatment planning algorithms.

2. Materials and methods

2.1. MCI method

To obtain the optimal dose distribution there are four distinct steps in MCI: definition of the geometry, backward transport, forward transport and optimization. A schematic presentation of the main MCI steps is shown in figure 1.

In the initial stage, a CT scan of the patient is used to define the geometry, described as a rectangular lattice of voxels. Material composition and density is defined according to the CT values, while tissue types (tumour tissue, critical tissue, normal tissue) are defined manually in a similar manner to conventional radiotherapy planning. The densities from the CT scans, and the PTV and critical structure contours are obtained from the treatment planning system. Only two materials are used in the geometry specification (water for densities \( < 1.1 \text{ g cm}^{-3} \), otherwise bone), although extension to more materials would be trivial, if a need was established. In the described version of MCI, the geometry is specified in two dimensions and the geometry in the third dimension is assumed to be invariant. Inclusion of the third dimension into the code would increase the overall computational time and storage requirements. This extension is the subject of ongoing work.

To obtain an initial guess of the intensity distribution, backward transport from the target volume (tumour) is performed first. The approximation of a uniformly distributed isotropic source within the target volume is used. Such a uniformly distributed source approximation leads to the intensity distribution that delivers a dose distribution which is too low at the edges of the target volume (see figure 2(b)). An algorithm similar to the one described in Liu et al (1993), might improve the initial guess, but the question of how to incorporate the requirements of the objective function in the initial guess remains. The energy distribution of photons is assumed to be the same as the energy distribution of an accelerator photon beam (without scattered contribution). From this source, backward or adjoint transport is performed. Particles exiting the patient geometry are scored on a cylindrical surface with the radius of 100 cm from the centre of mass of the target volume and grouped into a predefined number of equidistant and angularly divided bins, corresponding to the positions and directions at which particles start in the forward transport stage (possible gantry angles and collimator positions). Adjoint photon transport is used in order to simulate backward transport faithfully. In the adjoint transport the transport of the particles is inverted (i.e. instead of losing energy, particles gain energy in the scattering interactions). One of the reasons for the adjoint transport is the possibility for sensitivity analysis of the dose distribution within a patient to the phase space of the radiation
Figure 1. (a) Backward (adjoint) transport of photons from the target volume. These photons are transported out to the specified gantry and collimator positions, building up the intensity distribution used for the first optimization iteration. (b) In the forward transport stage, the dose for each bixel is calculated. A weighted sum (determined by the optimization) of the dose from each bixel in each field yields the dose distribution.

beams (for a more thorough discussion on adjoint transport in radiotherapy see Difilippo (1998)). Currently, adjoint transport is used only for determination of the initial intensity distribution, for which the advantage of adjoint transport over, for example, backprojection, has not yet been established. However, there is potential to couple the forward and backward transport stages for faster convergence to the optimal solution. It should also be noted that this step requires just a small fraction (a few per cent) of the total calculation time. Backward transport calculation is performed using the MCNP (Briesmeister 1997) Monte Carlo code, which has the capability of adjoint transport built into the code. The default 12 (energy) group photon libraries are used. Electron transport is at present not taken into account. A good review on adjoint transport can be found in Bell and Glasstone (1970).

Once the initial guess of the intensity distribution is obtained, the dose distributions of all beam elements (bixels), corresponding to narrow beams which make up a broad beam for each gantry angle, are calculated (forward Monte Carlo transport). Currently, a simple description of the narrow beam source is used. A narrow beam is represented as a point source at 100 cm from the isocentre, with the energy distribution modified from the 6 MV spectra of Mohan et al.
Monte Carlo-based inverse treatment planning

Monte Carlo simulation of the accelerator head should be performed and the obtained intensity distribution used. Information from the backward transport stage allows calculation of only those narrow beams that contribute dose to the tumour tissue. Less important beam orientations (with less weighting after the backward stage) are calculated less precisely to speed up the Monte Carlo calculation, because this is the most time-consuming phase of MCI. As the narrow beams are pre-calculated, no on-the-fly dose calculations are necessary during the optimization procedure. For the forward transport stage, the EGS4 Monte Carlo code (Nelson et al 1985) is used, because of the limitation of MCNP in the case of a large number of scoring voxels (Jeraj et al 1999). The usercode employed was an adaptation of RTPCART (Murray 1990) which transports particles in a Cartesian geometry. The precision of the Monte Carlo dose calculation (in a voxel in the middle of a bixel) varied from 1% in the theoretical examples to 0.2% in the clinical and 0.1% in the experimental example.

Using the information gained in the backward transport step, the initial guess of the intensity distribution for the optimization is calculated. To evaluate the treatment plan, at present a quadratic cost function based on dose criteria only, with different weights for different critical tissues and the target volume is used

$$OF = \sum_{T} \frac{C_{T}}{N_{T}} \sum_{i \in T} (d_{i} - D_{T})^{2}$$

where OF is the objective function, $C_{T}$ is the weight of a tissue type $T$ and $d_{i}$ is the dose in a voxel $i$ of type $T$. $D_{T}$ is a desired dose in a tissue type $T$ (required mean dose in tumour and 0 for other tissue types) and $N_{T}$ is the number of voxels of type $T$. In addition to so-called soft constraints (quadratic penalties), hard constraints were also defined, which limit the maximum dose to all structures and minimum/maximum dose to the tumour. The default values for $C_{T}$s are 1, although the choice of the weights is arbitrary. It was noted, for example, that higher values for tumour help the convergence. The $D_{T}$ values used are taken from Burman et al (1991). In general, choice of the parameters affects the final result and plan very much, and care should be taken as to how they are defined. Though the objective function is based on dose alone, incorporating radiobiological properties such as volume effects and/or radiobiological sensitivity is possible. Further investigation of the objective functions is outside the scope of this paper and not the main purpose of MCI. By stochastically varying the weights of the narrow beams, fast simulated annealing optimization (Szu and Hartley 1987) is used to determine the dose distribution that best satisfies the criteria specified by the objective function. The fast simulated annealing was written in-house and uses the standard Cauchy generator for grain sizes. The cooling schedules for both the temperature and grain are inversely proportional to the iteration number. Convergence to the optimum was tested by running optimization for different initial dose distributions. For quadratic objective functions, the method proved to be robust and convergence to the optimum was very insensitive to the choice of the optimization parameters (initial temperature and grain size, temperature and grain variation rate). When more complex objective functions (i.e. exponential for tumour, quadratic for normal tissue) were tested, selection of the parameters became much more important. However, optimization of the simulated annealing optimization schedule is a separate problem closely correlated to the objective function used and thus outside the scope of this paper. By using the simulated annealing optimization the intensity distribution for the ‘optimum’ treatment is achieved. Note that MCI only optimizes the intensities, not the positions of the beams.

MCI, developed so far, has been used for photon beams only. If other particles were used, some modifications to the described scheme would occur (e.g. definition of the optimal and
feasible set of energies in case of charged particles, definition of the source for the backward transport).

2.2. Examples

Several theoretical and clinical examples were calculated with MCI. One of the theoretical examples was simulation of a C-shaped target within a square body and a critical structure in the vicinity of the target. The $31 \times 31 \times 10$ cm$^3$ body was composed of 1 cm$^3$ water voxels (see figure 2(a)). There were 36 possible beam positions with 20 allowed orientations at each position. The bixel size was $1 \times 1$ cm$^2$ at the isocentre (100 cm from the source position).

As a clinical example, dose distributions for a lung tumour patient with a complex tumour shape were examined. The geometry was taken from CT data and determined on a $0.5 \times 0.5 \times 10$ cm$^3$ grid (see figure 3(a)). 21 beam positions and 21 bixel orientations at each position were allowed in the calculation. The bixel size was $1 \times 1$ cm$^2$ at the isocentre.
Finally experiment on an anthropomorphic phantom was performed to test the MCI method. A CT scan of the phantom was taken first. On the Cadplan\textsuperscript{TM} treatment planning system, the tumour shape was outlined as well as a critical structure (spinal cord). The shape of the tumour was an actual tumour contour for a patient, but was shrunk to completely cover the detector film (see figure 4). The geometry data were then passed to the MCI code, which produced a set of bixel intensities to be delivered by an accelerator. These parameters (gantry angle, jaw position, intensity) were then entered into the Lantis\textsuperscript{TM} accelerator interface and delivered by the Siemens Mevatron\textsuperscript{TM} accelerator. Seven equidistant beam positions (21 bixels within each field) were specified for the experiment. Gafchromic film, placed between two slices of the phantom with its centre at the isocentre, was used as a detector. The resultant
dose distribution recorded by the film was scanned and compared with the dose distribution calculated by MCI. The results were normalized to have the mean dose in the tumour equal to 100% in both cases.

3. Results

The dose distribution of the intensity distribution obtained after the backward transport stage for the **C-shaped target** phantom is shown in figure 2(b). It may be noted that this distribution approximately follows the target geometry, but is much less uniform than the optimized dose distribution shown in figure 2(c). The optimal dose distribution is symmetrical, as expected for a symmetrical geometry, which is an indication that the solution is close to optimal for the given objective function. Conformation of the dose to the target shape is good as well as sparing of the critical structure. The dose is also very uniform in the target volume, as requested.

Dose coverage of the tumour tissue for the **complex lung tumour** example was held within the ICRU limits of 95–107% (ICRU 1993). The dose distribution results using two different objective functions are presented in figures 3(b) and 3(c). In the example presented in figure 3(b), only the spinal cord was declared as a critical organ, where the dose should not exceed 60% of the mean dose to the tumour. In figure 3(c), the left lung was also defined as a critical organ, and a penalty was imposed for doses over 40% of the maximum. 21 fields were chosen to get good conformation to the tumour shape. Dose volume histograms for both cases are shown in figure 5.

The requested dose limits were completely fulfilled only in the first case with the spinal cord declared as the only critical organ. When the left lung was included as a critical organ (prescribed upper limit) all requested dose limits could not be met simultaneously. In this case the restrictions were too severe and as such had no physical solution. If the dose volume histograms are compared, it can be seen that strong sparing of a large area (i.e. left lung) inevitably leads to increased dose in other areas. Also dose to the tumour was much less homogeneous when the whole left lung area was spared, but, the target dose was still within the required limits (95–107%). As clearly seen, sparing of the lung increases the dose in the other critical organ—the spinal cord—and pushes it to the set limit (60%). It should not be forgotten that the detailed dose distribution (and consequently dose volume histogram) strongly depends on the chosen objective function and the weights of the variables used in the objective function.
Monte Carlo-based inverse treatment planning

Figure 5. Dose volume histograms (DVH) of two optimized plans. One plan (dashed line) was optimized with just spinal cord declared as a critical organ, where the dose should not exceed 60% of the mean dose to the tumour and one plan (solid curve), where in addition left lung was requested to have dose below 40% of the mean dose to the target. Note that in both cases dose to the tumour was held between the recommended ICRU limits (95–107%). Right lung DVHs were very similar for both cases and are thus not presented.

The typical calculation time for a two-dimensional plan with 1% precision in dose delivered to the tumour on $5 \times 5 \text{mm}^2$ resolution was a few hours on an average personal computer (Pentium II, 300 MHz). It should be noted that the time depends linearly on the number of narrow beams calculated (positions and orientations), and quadratically on the requested precision. For the example shown in figure 3, approximately 10 min were spent on the adjoint transport, approximately 4 h on the Monte Carlo dose calculation (with the average of $10^5$ particles per bixel) and approximately 1 h on the simulated annealing. MCI can be run on any computer platform that is supported by the Monte Carlo programs used (Unix, PC).

As a test of the reliability of MCI an experiment on the anthropomorphic phantom was performed. The geometry and experimental set-up of the film is shown in figure 4. Comparison between the experimental result and the result of the simulation is presented in figure 6.

As seen from figure 6, the agreement is very good, except in the central region, where the assumption of invariant geometry in the third dimension is not completely valid (bronchial region). The overestimation of the calculated dose in this region is up to 5%. By inclusion of the third dimension into the code, this discrepancy should be reduced. Overall discrepancies are random with a standard deviation of 3%. They are mostly due to statistical noise of the experimental read-out, which can also be seen from figure 6, because no systematic error is apparent.

4. Conclusions

MCI (Monte Carlo based inverse treatment planning), a new algorithm for inverse treatment planning based on Monte Carlo transport and simulated annealing optimization, has been developed. Monte Carlo transport is potentially the most accurate dose calculation algorithm for radiotherapy and ensures that the error due to approximate dose calculation is as small as possible. Simulated annealing optimization ensures that the optimized answer is close to the global minimum of the used objective function, which may be complex. Also, different
objective functions are easily incorporated into the simulated annealing algorithm. Though the optimal dose distribution has been calculated according to a simple dose-based objective function, the choice of the objective function, the constants defined within, and the coupling between the critical and target tissues are arbitrary.

There are four main steps in MCI: definition of the geometry, backward transport, forward transport and optimization. The geometry is specified on the basis of a CT scan of the patient and is at the moment assumed to be two-dimensional with the third dimension modelled as geometrically invariant. Backward transport is used to obtain a good initial guess of the intensity distribution. Adjoint Monte Carlo transport is used for this purpose. In the forward transport stage, the dose distribution of all bixels, which correspond to small beam elements of the multileaf collimator, is calculated by Monte Carlo transport. The optimization stage uses simulated annealing to minimize the specified objective function.

MCI plans have been generated for both phantoms and patient geometries. Three examples are presented in the paper. In the theoretical example of a C-shaped target, which has mirror symmetry, MCI gives an optimized dose distribution which is symmetrical as well, which is indication that the solution is close to optimum. The ability to obtain good conformation and uniformity of the dose to the target volume was demonstrated. Similarly, conformation of the dose to the tumour was achieved on a clinical example of a lung tumour. For comparison, a plan with explicit sparing of one lung is shown. Even for this case, the dose uniformity was within ICRU recommended limits in the tumour and the critical structure was spared. The final test of MCI was a dose comparison with experimental results from an anthropomorphic phantom. Complete treatment planning of a lung tumour was performed. The agreement between radiochromic film measurements and MCI predictions was within 3%, except in the bronchial region (up to 5% deviations), where the assumption of a geometrically invariant third dimension is not completely valid.

To conclude, in all cases examined, MCI plans exhibit proper functioning of MCI—good confirmation of the dose to the target volume and the avoidance of critical structures, as required by the objective function used. Comparison with the experiment showed that the calculated
Monte Carlo-based inverse treatment planning 1895

dose distribution matches the dose distribution given to the phantom/patient. A drawback of MCI is the relatively long computational time (a few hours for a two-dimensional plan with 5 × 5 mm² resolution and 1% statistical error), but with the increased speed of computers this will become less and less of a problem. Other limitations are that the geometry is currently specified in two dimensions only, and a point source is assumed for the incident photon beam. Work to overcome these limitations has been pursued. MCI has been applied to photon beams only, but with some modifications it could be applied to other particle beams as well. 3D Monte Carlo-based inverse treatment planning with accurate source modelling and simulated annealing optimization could be used as a benchmark with which to test the accuracy of other algorithms used for inverse treatment planning.

Acknowledgments

The authors wish to thank Wayne Beckham and Peter Hoban for valuable ideas, Martin Ebert for reading the manuscript and Martin Butson, Peter Metcalfe, Robin Hill, Mario Perez and Mark West for experimental help. PJK acknowledges the financial support of NIH grant CA 74158-02.

References

Bortfeld T 1997 Inverse treatment planning ESTRO Workshop on Challenges in Conformal Radiotherapy (Nice) (European Society for Therapeutic Radiology and Oncology)
Bortfeld T, Schlegel W and Rein B 1993 Decomposition of pencil beam kernels for fast dose calculation in three-dimensional treatment planning Med. Phys. 20 311–18
Hartmann-Siantar C L et al 1997 Validation and clinical implementation of the PEREGRINE Monte Carlo system for photon beam therapy (abstract) Proc. World Congress on Medical Physics and Biomedical Engineering (Nice) (Organizing Committee of the World Congress on Medical Physics and Biomedical Engineering) p 1102
ICRU 1993 Prescribing, recording and reporting photon beam therapy ICRU Report 50 (Washington, DC: ICRU)


Mohan R and Wang X 1996, Reply to physical vs. biological objectives for treatment plan optimisation Radiother. Oncol. 40 186–7


Nelson W R, Hirayama H and Rogers D W O 1985 The EGS4 code system Stanford Linear Accelerator Publication 265


Siebers J, Libby B and Mohan R 1998 Trust, but verify: comparison of MCNP and BEAM Monte Carlo codes for generation of phase space distributions for a Varian 2100C Med. Phys. 25 A143


