NOWOTWORY Journal of Oncology • 2004 • volume 54
Number 6 • 535–546

Invited article

Multiobjective inverse planning for external beam radiotherapy: decoupling the optimization and decision processes

Michael Lahanas[1], Eduard Schreibmann[1, 2], Lei Xing[2], Dimos Baltas[1] and Richard F Mould[1]

Inverse planning in radiation therapy is a trial and error process and many studies have been published which consider different algorithms, constraints, cases and objective functions. These planning algorithms combine in a specific way the objectives that are in conflict but do not provide the information that is necessary to obtain an optimal solution. Only in the last few years have significant increases in the availability of computing power enabled inverse planning to be performed to provide the information necessary to understand the possibilities of all dose distributions that can be obtained. We consider in this paper this multiobjective approach for external beam radiation therapy inverse planning that decouples the optimization and decision making processes. Inverse planning can now consider the number of beams, their orientation and optimal beam fluences and their dependence on importance factors. In this way it is possible to exploit the possibilities of advanced technologies such as intensity modulated radiation therapy (IMRT) or tomotherapy. Using data mining and visualization techniques a solution can be selected that requires, if possible, a small number of beams, and that also provides the required dose delivery to the target. Additionally, it allows reduction of the dose in the healthy tissue, especially in organs at risk, in such a way that the compromise we have to make for all the objectives in comparison to the best individual values is as small as possible.

Planowanie leczenia typu „inverse planning" z zastosowaniem wiązek zewnętrznych – rozdzielenie procesów optymalizacji i podejmowania decyzji

Planowanie z zastosowaniem techniki „inverse planning" w radioterapii oparte jest na metodzie prób i błędów. Opublikowano wiele badań przedstawiających różne algorytmy, ograniczenia oraz przypadki z zakresu tej dziedziny. Algorytmy stosowane niegdyś w planowaniu typu „inverse planning" uwzględniały złożony zestaw założeń, które, choć powiązane, pozostawały jednakże w sprzeczności, a jednocześnie nie dostarczały danych niezbędnych dla uzyskania optymalnych rozwiązań. Znamienny rozwój technik komputerowych, obserwowany w ostatnich latach, stworzył możliwość prowadzenia tego typu planowania, z pełnym wykorzystaniem możliwych do uzyskania sposobów dystrybucji dawek.

W niniejszej pracy przedstawiamy planowanie leczenia typu „inverse planning" z zastosowaniem wiązki zewnętrznej, umożliwiające oddzielenie optymalizacji od procesu podejmowania decyzji. W chwili obecnej planowanie takie pozwala uwzględnić liczbę wiązek, ich kierunek oraz optymalną penetrację. Dzięki temu możliwe stało się pełne wykorzystanie zalet zaawansowanych technik, takich jak IMRT lub tomoterapia. Dzięki dokładnej analizie danych i zastosowań metod wizualizacji można opracować rozwiązania wymagające stosowania niewielu wiązek, a jednocześnie zapewnić prawidłowe rozmieszczenie dawek. Równocześnie można do minimum ograniczyć dawkę skierowaną na zdrowe tkanki, co ma szczególne znaczenie w przypadku narządów kluczowych. Dzięki temu łatwo jest pogodzić wszystkie założenia skutecznej i bezpiecznej radioterapii, co pozwala ograniczyć ryzyko powikłań.

Słowa kluczowe: radioterapia, planowanie typu „inverse planning", techniki komputerowe
Key words: radiotherapy, inverse planning, computer technology

Introduction

In the year 2003, worldwide, some 8 million persons were diagnosed with cancer and it is estimated that this number will increase to 20 million by the year 2020. Currently at least 50% of cancer patients are treated with radiation therapy in the USA [1] and 67% in the European Union. Technological advances enabling the provision of more controlled beam fluence distributions, such as with

[1] Department of Medical Physics & Engineering, Strahlenklinik, Klinikum Offenbach, Starkenburgring, Germany
[2] Department of Radiation Oncology, Stanford University School of Medicine, USA

intensity modulated radiation therapy (IMRT), enables the achievement of dose distributions of higher quality than was previously possible using only filters and wedges.

The objectives of radiation therapy are to produce a treatment plan which meets target radiation doses for the Planning Target Volume (PTV) which includes as well as the Gross Tumour Volume (GTV) an additional margin to take into account, for example, positional inaccuracies and patient movements. Additionally the objectives include minimizing the damage to nearby structures such as the surrounding normal tissue (NT) and organs at risk (OARs).

Dose is the amount of energy deposited per unit of mass. The physical and biological characteristics of the patient anatomy and of the source, such as intensity and geometry, are used for the calculation of the dose function, *i.e.* the absorbed dose at a point in the treatment volume. The dose distribution specifies the corresponding three-dimensional non-negative scalar field. A physician prescribes the required dose function *i.e.* the absorbed dose, as a function of its location in the body.

The treatment plan is expressed as a set of objectives that due to physical limitations cannot be always satisfied. This is because the objectives are sometimes in conflict and we thus have a multiobjective or multi-criteria problem to solve.

The radiation oncologist uses for the evaluation of the dose distribution quality a cumulative dose volume histogram (DVH) for each structure (PTV, NT, OARs). This displays the volume fraction that receives at least a specified dose level. The objectives are called DVH-based objectives if expressed in terms of DVHs related values.

In external beam radiotherapy, high energy photon beams are emitted from a source on a rotating gantry with the patient positioned according to treatment plan requirements, which are sometimes that the tumour is at the centre of the rotation axis. The beam can be shaped in various ways as it rotates around the patient and the radiotherapy treatment plan specifies the beams' shapes and intensities at a fixed number of source angles. The beam shape can be modelled by using blocks, wedges and other devices.

Three-dimensional conformal radiotherapy (3D-CRT) is an advanced form of external beam radiation therapy in which the high-dose treated volume is planned to encompass the PTV, at the same time minimizing the dose to the surrounding OARs. It requires the use of three-dimensional imaging methods and is typically accomplished with a set of fixed radiation beams shaped according to the projections of the target. The radiation beams typically have a uniform intensity across the field, or the intensity is modified by devices such as wedges or compensating filters.

A major quantity of interest is the dose $D(r) = D(x, y, z)$ at a point r in the treatment volume. The dose distribution specifies the corresponding three-dimensional non-negative scalar field $D : R^3 \rightarrow R^+$, $\mathbf{r} \rightarrow D$. The dependence between energy fluence Ψ and dose distribution D, $A \cdot \Psi = D$, is given by the energy absorption per mass and energy fluence unit operator A or *energy absorption operator* which describes the local energy dissipation for a given patient anatomy. A dose distribution is possible if there is a source distribution which is able to generate the distribution.

The *dose space* $\{D\}$ defines the space of all achievable dose distributions for a given patient anatomy and consequently the set of all possible energy fluences defines the fluence space $\{\Psi\}$. The physical and biological characteristics of the patient anatomy and of the source are used for the calculation of the *dose function*, *i.e.* the absorbed dose as a function of the location in the body. A physician, depending on the patient, prescribes the required dose function.

While the determination of D from Ψ, the solution of the so-called *forward problem* is possible, the *inverse problem*, *i.e.* determination of Ψ for a given D is not always possible. The forward problem is the dose calculation problem for which a unique solution exists. Even if A can in principle be inverted, the image of $A^{-1}D$ will not be always an element of $\{\Psi\}$ and it is possible that the solution is not unique. This problem can be solved analytically only for very simple cases. As an analytical solution cannot be obtained we consider the inverse problem to determine the position and number of beams, the beam weights, fluence distribution, etc, such that the obtained dose distribution satisfies as nearly as possible the required dose function via an optimization process. This process is called *inverse optimization* or *inverse planning*.

For the term optimum we have to consider the constraints and limitations that we use. It is common to select the number of beams and their orientation from experience gained by treatments with similar cases. Then the inverse planning can be limited to finding the optimal fluence distribution for the specific beam topology. It is obvious that the optimum solution found is not necessary a global optimum solution as other better solutions could be obtained for other beam configurations. Additional other limitations are the use of only a limited number of possible candidate solutions with beams at specific angles.

Radiotherapy treatment planning in the 21st century is unfortunately still a trial and error process. The planner has a list of dose values and fraction of volumes for each OAR that should not be exceeded. The problem is that radiation from outside the patient has to pass partly through the body and through OARs and has to deliver a specific dose level to the tumour. The protection of OARs and the coverage are objectives that are in conflict. One possibility to reduce the dose in the OARs is to increase the number of beams that allows a reduction in the intensity of each beam, but this also increases the complexity of the treatment plan.

The determination of an optimal number of beams and their orientation is important as a reduction of the number of beams simplifies the treatment plan in terms of time and complexity. Consequently, it also reduces the possibility of treatment errors.

There are many different ways to protect a specific OAR but protecting all OARs simultaneously in an optimal way is not possible. The planner, does though, have some possibilities with current treatment planning systems to modify the dose distribution. A parameter which is the *weight* or *importance factor* can be assigned to each objective representing to represent its importance in achieving for the objective an optimal value. Additional critical dose values and volume fractions can be set.

Based on experience, a specific beam setup (a number of beams and their orientations) is selected and an algorithm is used to optimize the beam fluence distribution for the specific set of weights and parameters selected by the planner. This so-called *inverse optimization of beam fluences* is fast enough for clinical practice: requiring from only a few seconds to a few minutes. The optimization result usually does not satisfy some of the objectives and the treatment planner by modifying the parameters has to repeat the optimization. In a trial and error process, by comparison with the previous result the search is continued until the planner considers that it is difficult to improve further the resulting dose distribution or that the dose distribution obtained is acceptable.

Multiobjective inverse planning

Inverse planning methods which are currently in use assume that the planner has an *a priori* knowledge of the trade-off between the objectives and their limitation, expressed by specific weights and critical dose values. The decision process precedes the optimization. However, that such knowledge does not exist is seen by the fact that the optimization is repeated multiple times. Because of this, inverse planning can be described as an iterative process and the optimization process that continues until a solution is selected. However, for this solution we do not know its actual quality except that it is superior to the previous examined solutions; see Figure 1(a). It must therefore be realised that in actual fact, only pairs of solutions are compared.

Xing *et al* [2] proposed a method that searches using an optimal set of importance factors and provides a solution based on the maximization of a utility function derived from dose-volume histograms. This is an approach that simulates the *trial and error* approach of treatment planning. The treatment planning decision is described by a utility function.

A similar method was suggested by Yan Yu [3] who used artificial intelligence that specifies goals which if they cannot be satisfied can be relaxed. Other methods have also been proposed to determined optimal importance factors. [4] However, all these methods do *not* provide trade-off information, and whether the solution is in fact an optimum is in the final analysis unknown to the treatment planner. This situation is therefore not ideal.

Only in the last few years have inverse planning methods been considered that decouple the decision and the optimization processes. Coupling is a consequence of the trade-off between the objectives in conflict. Also, a consequence of the trade-off is that a single optimal solution does *not* exist but a *set* of optimal solutions that require a decision process as to which of this *set* of solutions finally should be used. We have therefore to consider a multiobjective optimization problem.

It is emphasised that for multiobjective problems it is usually not possible to simultaneously satisfy all objectives in an optimally possible way with the same set of parameters. Each of the objectives is represented by some function that is used to evaluate the quality of a solution for the specific objective, i.e. a score function or objective function. We do not have a single objective function but a *set* of functions, a multi-valued function for which we have to find a set of parameters such as beam weights, orientations and wedge angles which optimizes all individual objective functions.

Mathematicians considered 'What is an optimum?' in the case of a multi-valued function almost at the same time when radiotherapy was first used at the end of the 19th century. In particular, Vilfredo Pareto [5] extended the definition of optimality for multi-valued functions used in economics.

For multiobjective problems it is more likely that a single function does not exist but rather a finite set or even an infinite set of optima. The so-called *Pareto optimal solutions or non-dominated solutions* forming the *Pareto set*. [6, 7] With the advances in computing power and progress in multiobjective evolutionary and deterministic optimization algorithms, more problems can now be considered with *real* multiobjective optimization algorithms and provide not only better solutions but also an insight to the nature of the problem itself which is under consideration.

The purpose of multiobjective optimization is to obtain a representative set of the entire Pareto set. This set is used to analyze the trade-offs between the objectives in conflict and then to select a solution that satisfies simultaneously, at best, all objectives. This is the decision making process performed after the optimization. It is not necessary to repeat the optimization. We say that the optimization precedes the decision making process and that optimization and decision are *decoupled*; see Figure 1(b).

Why do we not try to find automatically the single solution that is finally selected? The problem is that a uniquely recognized utility function does not exist that characterizes the optimality of a solution and more importantly the selection requires knowledge that is obtained only *after* the representative set is known and analysed.

The problem is thus transformed from an optimization problem to a decision making problem. Even if we had such a utility function we do not know how to find this solution, i.e. what parameters and weights or even objective functions should be used to find the optimal beam fluences. The problem is that treatment planning systems try to be as simple as possible and having a large number of solutions is something that

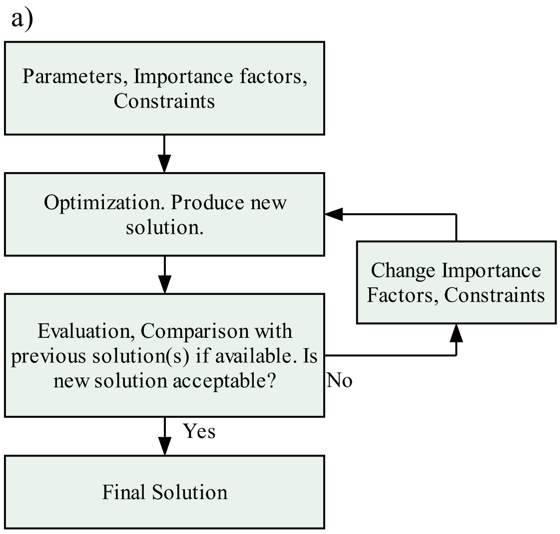

Figure 1 (a). Traditional inverse planning
(b) Multiobjective inverse planning that decouples the optimization and decision process

makes the planning more difficult. There is a considerable amount of work which has taken place in trying to simplify the planning procedure by methods that search for a single optimal solution. Even with many planning methods proposed using different set of objective functions the trial and error method still is that which is the most time consuming process in inverse planning.

The multiobjective optimization approach has some additional important benefits to those already described. It provides a coherent and complete view of all possible solutions that none of the current planning systems offer. It decouples the decision and optimization process and eliminates the guess work of the past that provides a solution of unknown quality. Only if we have a representative set of non-dominated solutions for each method, set of objectives and algorithms proposed, we will able to compare the results and understand the true physical dose limitations of a treatment method, see Figure 2.

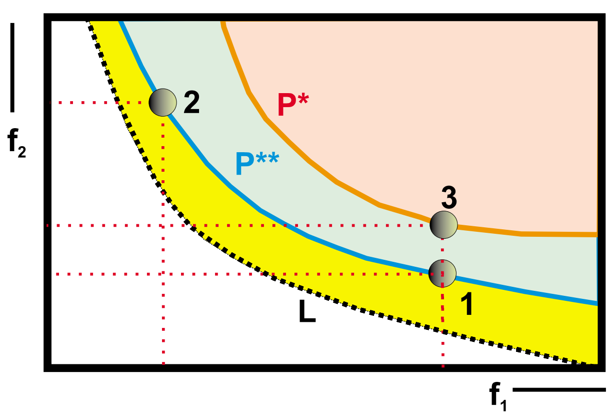

Figure 2. Example of a trade-off in radiotherapy. Trade-off due to physical limitations L and two other possible trade-offs P* and P** due to other limitations, such as number of beams, objective functions, machine resolutions etc. Some possible inverse planning results are shown; two non-dominated solutions 1, 2 and a dominated solution 3. The aim of multiobjective inverse planning is to compare different treatment methods and for each method to obtain the trade-off from a representative set of solutions that is used to select the best possible solution from this set

The comparison has to be made using a combination of objective values and dose-volume histogram related values derived from the non-dominated sets. The comparison has to consider trade-offs between these quantities as it is not enough to consider individual values separately.

The most important aspect of multiobjective inverse planning is that solutions can be obtained which provide the best possible protection for the patient in terms of irradiation of critical organs and tissues (OARs, NT) together with a simultaneous achievement of as maximum as possible coverage of the tumour region with the required dose.

Geometrical optimization of beam orientations

The aim of beam orientation optimization is to determine a configuration with a small as possible number of beams such that a desired dose distribution can more likely be achieved. A single beam would deposit a very high dose to the NT. Using more beams it is possible to increase the dose in the tumour, and retain the dose to the surrounding healthy tissue at a sufficient low level. However, this increases the treatment complexity.

The combinatorial complexity of this problem is extremely large and current treatment planning systems do *not* consider the beam orientation problem in inverse planning. Instead the number of beams and their orientation is specified by the planner based on the particular case being planned and on experience gained by the planner with other similar cases. Various theoretical studies have been presented on how the results depend on the number of beams and their orientation: but using only raw approximations [8].

If we have to choose M beams out of N possibilities then the number of possible beam configurations is $N!/M!(N-M)!$. For $M=3$ beams out of $N=72$ beams

59,640 combinations exist. The number of combinations for the range M=3-9 and N=72 is shown in Figure 3. The dependence is almost exponential on M. For M=9 and N=72 we have 8.5x10[10] combinations. An increase of the number of optimizations by a factor 100 or more is necessary if for each possible configuration a sufficiently large set of importance factors has to be scanned.

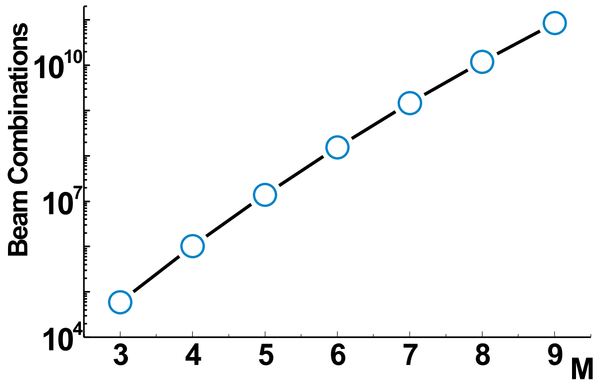

Figure 3. Number of combinations of M beams from 72 possible beams as a function of *M*

Many more combinations exist if non-coplanar beam configurations are considered. Even if many beam configurations obviously can be ignored the problem remains combinatorial complex and only a very small subset of configurations can be examined.

One of the first approaches for the beam orientation problem was to decouple the beam orientation optimization from the fluence distribution optimization. Haas *et al* [9-11] proposed the use of multiobjective evolutionary algorithms (MOEAs) to obtain the trade-off information necessary for the selection of an optimal beam configuration. Also, Haas *et al* [11] used cost functions derived from geometric considerations and the multiobjective algorithm NPGA [12] to obtain the non-dominated set for the geometric objectives.

Simplifications such a limitation of the method in two-dimensions using the most representative 2D computed tomography slice in the plan were used. Instead of the planning target volume (PTV) the planning target area (PTA) is considered. The geometric objective functions to be minimized are as follows.

• Difference between the area where all M beams overlap and the area of the PTA

$$C_{PTA} = area\left(B_1 \cap B_2 \ldots \cap B_{N_{beam}}\right) - area\left(PTA\right) \tag{1}$$

• Overlap area between each beam and the j[th] OAR

$$C_{OARj} = \sum_{i=1}^{N_{beam}} \rho_i \frac{area\left(B_i \cap OAR_j\right)}{N_{Beam}},$$
$$\rho_i = \begin{cases} \beta\left(\delta_{PTA} - \delta_{OAR}\right) & \delta_{OAR} < \delta_{PTA} \\ 1 & \delta_{OAR} \geq \delta_{PTA} \end{cases} \tag{2}$$

δ_{PTA} and δ_{OAR} are distances shown in Figure 4 and β is a parameter that favours beam entry points further away from OARs.

• Overlap from pair wise beam intersections to minimize hot spots

$$C_{NT} = \sum_{i=1}^{N_{beam}-1} \sum_{j=i+1}^{N_{beam}} area\left(B_i \cap B_j\right) \tag{3}$$

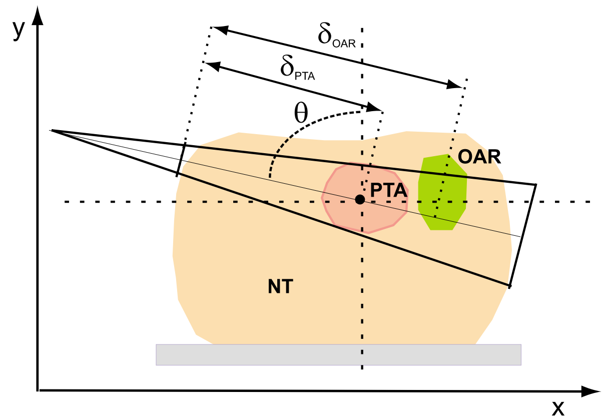

Figure 4. Geometric parameters used for the solution of the beam orientation problem. The gantry angle θ of a field (beam) is shown. The patient body including the normal tissue NT, one organ at risk (OAR) and the planning target area (PTA) which includes the tumour is shown

An integer representation of the beam gantry angle was used and the length of each chromosome is equal to the number of beams involved in the plan. A particular solution, i.e. a chromosome, is represented as a vector $C^T = (\theta_1, \ldots, \theta_{N_{beam}})$, where θ_i is the i[th] individual beam gantry angle.

An intermediate recombination operator is used [9], such that the parents C_{P1} and C_{P2} produce the offspring $C_o = round(C_{P1} = \gamma(C_{P2} - C_{P1})]$, where γ is a random number in the interval [-0.25 1.25]. A mutation operator is used to introduce new beam angles into the population by generating integers that lie in the range [0°, 359°].

It is also important to note the inclusion of problem specific operators which attempt to replicate the approach followed by experienced treatment planners. Such an operator is used to generate *k* equispaced beams as this distribution will reduce the area of overlap between the beams. One gantry angle from a particular chromosome is selected randomly positioning the *k*-1 remaining beams evenly. A further mutation operator is used to perform a local search by shifting randomly by a small amount (less than ±15°) one of the selected beam gantry angles.

Schreibmann et al [13] extended this approach to realistic three-dimensional treatment planning using a set of geometric objective functions *g* that correspond to dose variance-based objective functions used commonly in radiotherapy: for the PTV the dose variance f_{PTV} around the prescription dose D_{ref}^{PTV}, for NT the sum of the squared dose values f_{NT} and for each f_{OAR} the variance for dose values above a specific critical dose value D_{crit}^{OAR}. We use d_j^{PTV}, d_j^{NT} and d_j^{OAR} for the dose values at the j[th] sampling point for the PTV, the NT and each OAR respectively. N_{PTV}, N_{NT} and N_{OAR} are the corresponding number of sampling points. The geometric objective functions *g* and the corresponding dosimetric objective functions *f* are as follows.

• g_{PTV} for the coverage of the PTV. The geometrical formulation of this objective takes into account the relative volume of the PTV covered by the intersection of the beams which should be maximized. In order to use a minimization algorithm the following objective function is used.

$$g_{PTV} = 1 - \left[\frac{PTV \cap \bigcap_{i=1}^{N_B} B_i}{PTV} \right], \qquad f_{PTV} = \frac{1}{N_{PTV}} \sum_{j=1}^{N_{PTV}} \left(d_j^{PTV} - D_{ref}^{PTV} \right)^2 \tag{4}$$

$\bigcap_{i=1}^{N_B} B_i$ is the intersection volume of the N_B beams B_i, $i=1,2,\ldots,N_B$. This function is defined in an analogue way as the Conformal Index COIN [14].

• g_{NT} for the protection of the NT. This can be achieved by using spatially distributed beams. The objective is to minimize the overlap volume between any pair combination of beams:

$$g_{NT} = \frac{\sum_{i=1}^{N_B - 1} \sum_{j=i+1}^{N_B} \left(B_i \cap B_j \right)}{\sum_{i=1}^{N_B} B_i}, \quad f_{NT} = \frac{1}{N_{NT}} \sum_{j=1}^{N_{NT}} \left(d_j^{NT} \right)^2 \tag{5}$$

Beams having close entrance points will generate a large intersection volume in the overlap region and thus a high value for this objective. As the entrance points are more widely spaced, the intersection volume between the beams is minimized, ideally tending to the PTV volume.

• g_{OAR}^k for the protection of the k[th] OAR. The objective is to minimize the exposure of the OARs formulated as minimization of the intersection volume between the beams and the OAR:

$$g_{OAR}^k = \sum_{i=1}^{N_B} \frac{\left(OAR_k \cap B_i \right)}{OAR_k}, \quad f_{OAR}^k = \frac{1}{N_{OAR}^k} \sum_{j=1}^{M} \left(d_j^{OAR} \right)^2, \qquad k=1,2,\ldots,M \tag{6}$$

B_i, $i=1,2,\ldots,N_B$ are the N_B beams in the plan and *M* is the number of OARs to be considered. Ideally, there is no intersection between the OAR and any of the beams, and thus the fraction is zero.

The terms $(OAR_k \cap B_i)/OAR_k$ are pre-calculated and stored for all possible beams before the optimization. The calculation of these terms during the optimization would considerably increase the optimization time. The final cost function g_{Tot} for the geometrical optimization is a sum of the above objectives:

$$g_{Tot} = w_{PTV}\, g_{PTV} + w_{NT}\, g_{NT} + \sum_{k=1}^{M} w_{OAR}^k g_{OAR}^k \tag{7}$$

where, w_{NT}, w_{PTV}, w_{OAR}^k, are the importance factors for the objectives of the NT, PTV and the k[th] OAR.

A multiobjective optimization provides the trade-off information for the geometric formulated objectives. Schreibmann *et al* [13] have shown that there is a correlation between the geometric and dosimetric objectives and this permits finding beam configurations that are more likely to produce high quality solutions.

Multiobjective fluence optimization

In IMRT, which is an advanced conformal radiotherapy method, each beam is divided in a number of small beamlets (bixels). The intensity of each beamlet can individually be adjusted. A dose matrix is precalculated and contains the dose value at each sampling point from each bixel with unit radiation intensity. The intensity (weight) of each beamlet has to be determined such that the produced dose distribution is *optimal*. The number of parameters can be as large as 10^4.

The first application of multiobjective optimization of beam fluences for IMRT was given by Cotrutz *et al* [15]. Later, Lahanas *et al* used a more efficient algorithm, L-BFGS [16], alone [17] or in combination with an evolutionary algorithm [18] for the optimization of the intensity distribution in IMRT where the orientation and the number of beams are fixed. As an illustration, depending on the number of OARs to be considered, we have 3-7 objectives. The results have been compared with another set of objective functions, see Eq. (8), which have been used by others.

$$f_{OAR} = \frac{1}{N_{OAR}} \sum_{j=1}^{N_{OAR}} \max\left(0, \left(d_j^{OAR} - D_{OAR}^C \right)^2 \right) \tag{8}$$

The function in Eq. (8) only considers the avoidance of dose values above a critical level D_{OAR}^C in the critical structures. The comparison of the non-dominated sets using Eq. (6) and Eq. (8) respectively have shown that Eq. (6) for is preferred to Eq. (8) as by using Eq. (6) f_{OAR} it is possible to suppress also smaller dose values than D_{OAR}^C that are not *visible* using Eq. (8).

Multiobjective optimization of number of beams: their orientation and fluence distribution

Most of the currently used methods for inverse planning in radiation therapy consider only the determination of the beam weights, intensity profiles or wedge angles for a specific configuration of beams and a specific set of importance factors. A much larger search space has to be considered if also the optimal numbers of beams and their orientations have to be found. Furthermore, results depend on the importance factors used.

Many possible combinations can be formed in the presence of multiple objectives, leading to an enormous number of possible solutions among which the best is to be found. The geometric optimization provides beam orientations that more likely provide good solutions in comparison to other random configurations. Geometric algorithms although, cannot take into account that the intensity distributions of beams are interdependent and the orientation of beams and fluence distribution cannot completely be decoupled.

Various recent studies considered the simultaneous optimization of the beam orientation and fluence optimization. Single objective algorithms were used using specific importance factors and fixed number of beams [19-24]. If the beam configuration is considered then the Pareto set is not necessary convex. The algorithms described also provide dominated solutions.

Conventional optimization algorithms search for an optimal solution by comparing the fitness values between two solutions. In the Pareto ranking based algorithm the dominance relation or, if constraints are used, the constrained dominance relation is used for the search. For different beam configurations the fitness is not a guarantee that only non-dominated solutions will be selected. Also optimal importance factors may depend on the number of beams and their orientation.

Schreibmann *et al* [25] recently applied for the first time multiobjective inverse planning where *no* importance factors are used. The user specifies for consideration a minimum and maximum number of beams, usually in the range 3-9. A hybrid evolutionary algorithm *NSGA-IIc* [26] was applied with the deterministic algorithm *L-BFGS* used for the optimization of the intensity distribution.

The role of the evolutionary component is the search in the beam-configuration space, leaving the role of intensity optimization to powerful gradient based optimization routines. The algorithm provides a representative set of efficient solutions. Clinically acceptable results can be obtained in only one hour.

More than 5000 archived solutions are obtained after 200 generations using a population size of 200 solutions. An arithmetic crossover is used with a random mixing parameter a in [0,1] and a flip mutation. A mutation and crossover probability 0.01 and 0.9 respectively are used. Three cases have been considered for testing the algorithm, a cube shaped PTV situated in the middle of a bigger cube; simulating the NT, see Figure 5(a), a prostate case and a head and neck case.

For a test case of a cube shaped PTV in a cubed NT the dosimetric variance f_{NT} and f_{PTV} trade-off is shown in Figure 5(b) as a function of the number of beams. The results show that the largest protection gain is obtained by changing from 3 to 4 beams. The dose homogeneity can be improved but the dose variance in the NT increases rapidly. There is no extra benefit by using more that 3 beams. Only angles at 0, 90, 180 and 270 degrees are dominant. All other beam directions have to travel a larger distance through the NT.

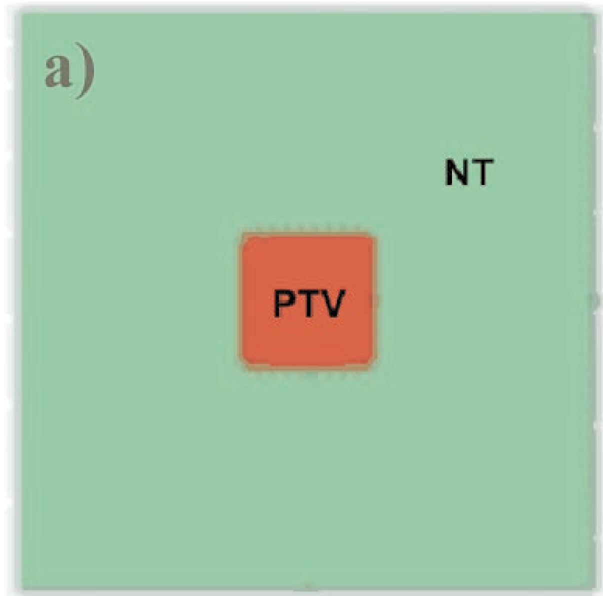

Figure 5a

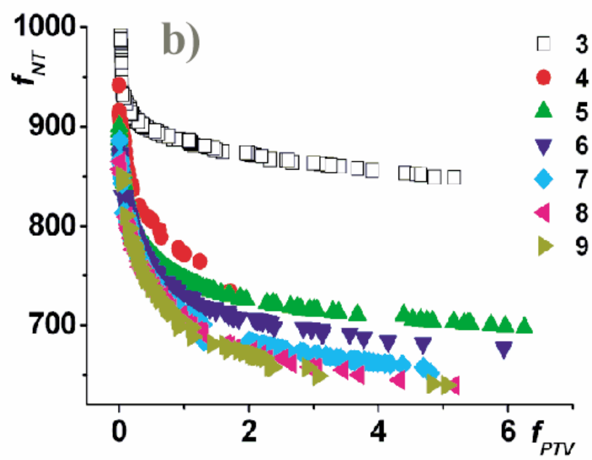

Figure 5b

Figure 5. (a) A phantom test case
(b) Trade-off $f_{PTV} - f_{NT}$ obtained by *NSGA-IIc* for this case.
The dependence on the number of beams is shown

For the prostate case, see Figure 6, only a few beam directions were dominant and necessary that can avoid passing though the OARs.

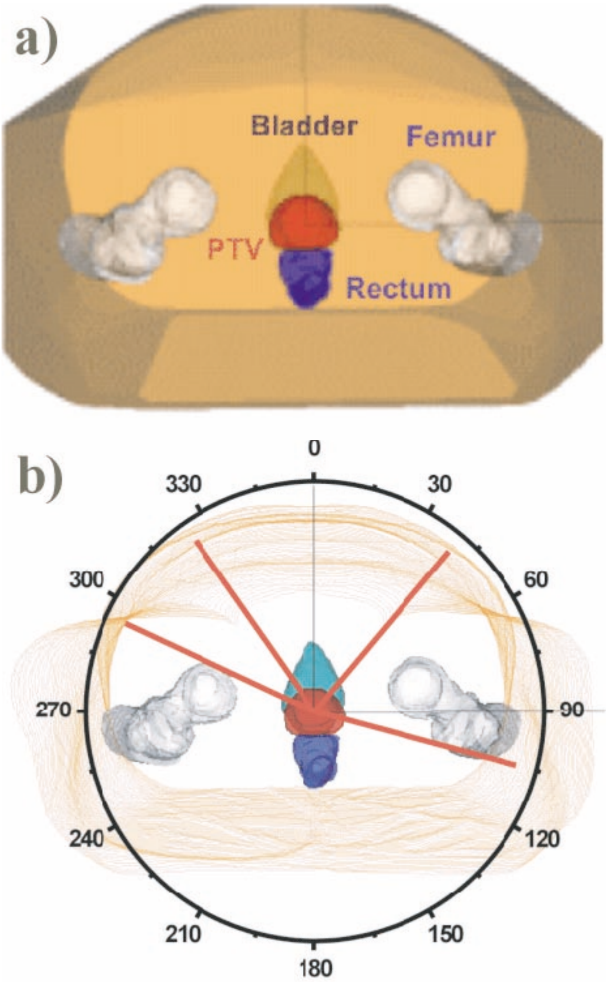

Figure 6. (a) The prostate case
(b) Gantry angles that are more likely selected by *NSGA-IIc* for this case

A difficult head and neck case, see Figure 7, was considered where the beam has to pass through OARs. Projections of the Pareto set show the trade-off between the objectives. For the spinal cord and the left and right parotids the dose variances f_{SC}, f_{ParL} and f_{ParR} were used as objectives.

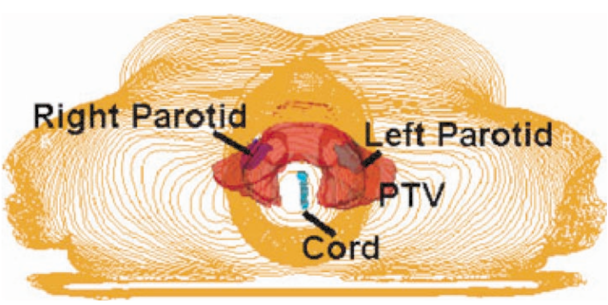

Figure 7. The head and neck tumour case

The $f_{NT} - f_{PTV}$ trade-off, Figure 8(a), shows that above 6 beams there is no significant decrease of the dose variance in the PTV. Also the variance in the normal

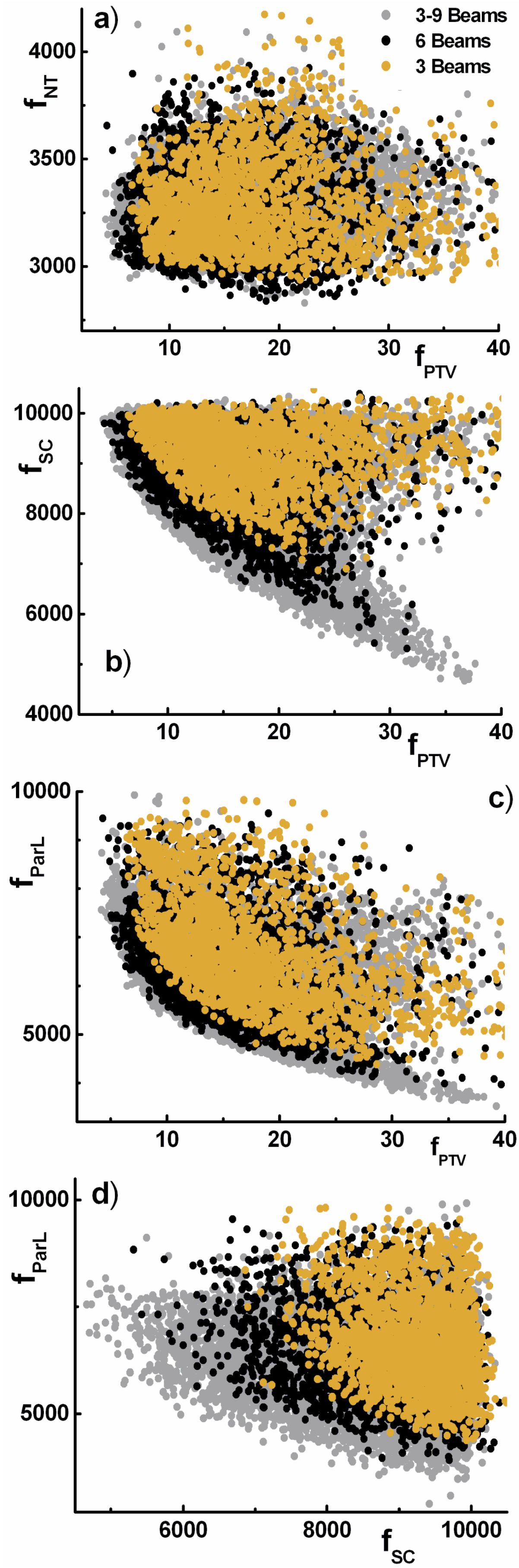

Figure 8. Trade-off between a) (f_{PTV}, f_{NT}), b) (f_{PTV}, f_{SC}) c) (f_{PTV}, f_{ParL}) and d) (f_{SC}, f_{ParL}) on the number of beams (3-9, 3-4 or 6) for the filtered solutions

tissue is not as in the two previous cases where the normal tissue can be protected better with more beams. Similar the $f_{PTV} - f_{SC}$ and $f_{PTV} - f_{ParL}$ correlations,

Figures 8(b) and 8(c) show that for more than 6 beams there is no significant improvement with increasing number of beams. The $f_{SC} - f_{ParL}$ trade-off, Figure 8(d), shows that the problem is that simultaneous objectives are protection of the spinal cord and of the parotid glands. The result is better the more beams we use and at least 6 beams are required. Additionally both spinal cord and the parotids cannot be optimized simultaneously. In combination with Figure 8(c) we see that the dose variance f_{PTV} which is correlated to the coverage of the PTV with the prescription dose requires that we have to accept a strong increase in the variance for f_{PTV} values below 10.

More important are some dosimetric parameters that show the limits of the dose distributions that can be obtained such as the average and minimum dose in the spinal cord and the minimum dose in left parotid as a percent of the prescription dose and the coverage of the PTV at 95% of D_{ref}.

The minimum and average doses in the spinal cord depend on the number of beams, see Figure 9. This dependence decreases as the PTV coverage increases and approaches a limit of 98% for more than 6 beams and a limit of 95-96% for 3-4 beams. To better protect the spinal cord a large as possible number of beams is required, but a much smaller coverage for the PTV must be accepted. For 85% coverage the minimum dose in the spinal cord can be reduced only to 60% of the prescription dose for the spinal cord.

This case shows the limitations of IMRT for this particular case. Such information provided by multiobjective inverse planning is not provided by any other method.

For the multiobjective problem it is important to apply visualization techniques to understand the multidimensional results. Figure 10(a) shows the percentage of spinal cord and brainstem that receive a dose above the critical level. A set of 13,462 solutions is shown obtained from inverse planning that includes beam orientation optimization using 6-12 beams.

In Figure 10(b) the PTV dose coverage of the solutions is shown, i.e. the DVH value for the PTV for a dose which is 95% of the prescription dose. The results now show that in order to obtain solutions with a good protection for the brainstem and the spinal cord simultaneously we have to accept a coverage value as low as 80% at 95% for D_{ref}. The main problem is the spinal cord. We can achieve a dose coverage value of 95% but the spinal cord cannot be protected as desired.

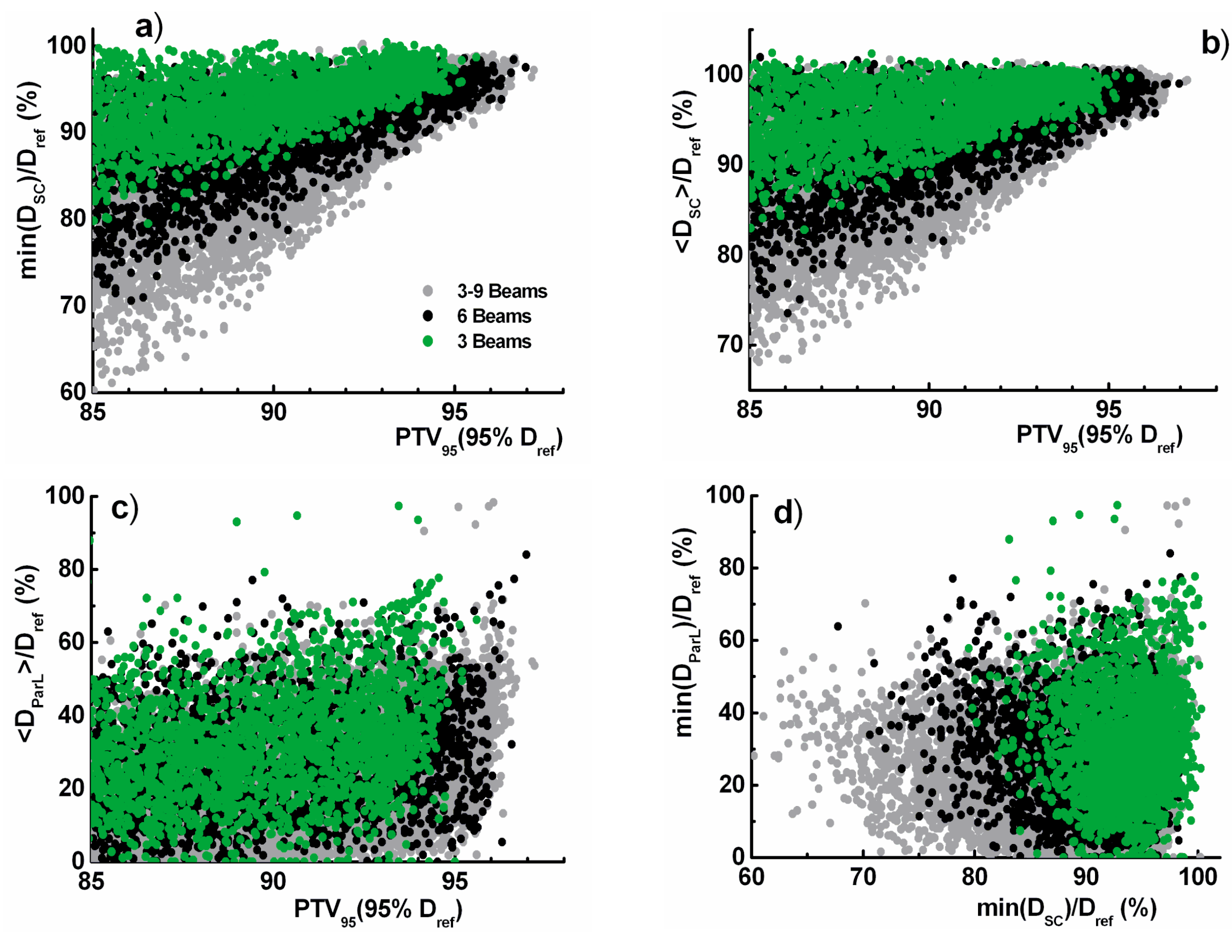

Figure 9. Dependence of trade-offs between dosimetric parameters for the head and neck case. (a) (PTV_{95}, min(DSC)) (b) (PTV_{95}, <DSC>) (c) (PTV_{95}, min(D_{ParL})) and (d) (min(DSC) – min(D_{ParL})) on the number of beams (3-9, 3-4 or 6) for the filtered solutions

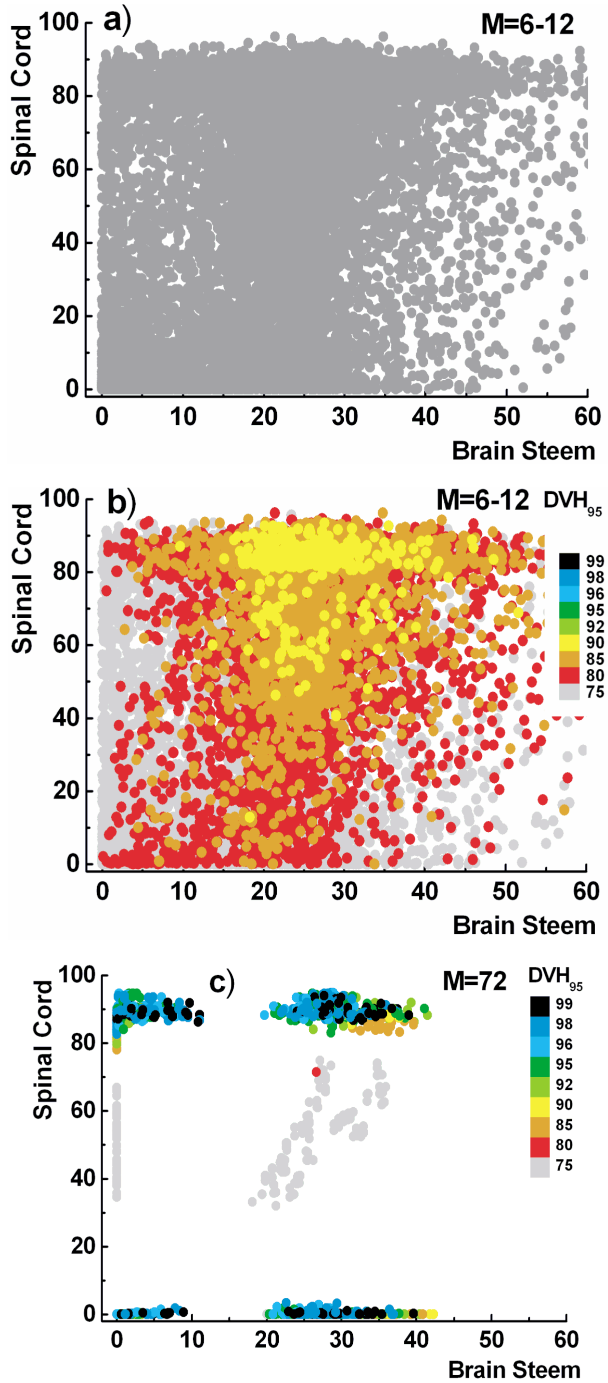

Figure 10. Distribution of the overdosage volume in the spinal cord and the brainstem for non-dominated solutions (a) for 6-12 beams (b) same as (a) but using different colours to represent the corresponding DVH for the PTV at 95% of the prescription dose (coverage) (c) result using 72 beams (rotation therapy). In this case we have solutions that can protect simultaneously the spinal cord and the brainstem achieving a large PTV coverage

In Figure 10(c) we show the result of 2,250 solutions using 72 beams, i.e. using a rotation therapy. The importance factors were scanned uniformly but the solutions are clustered, showing that there is a nonlinear complex mapping from importance to objective/dosimetric space. Now there is the possibility to obtain solutions with 95% coverage that protect both the spinal cord and the brainstem.

This case shows that it is not true that the dose distribution in IMRT cannot be significantly improved by using more than even 12 beams. For the head and neck case only, a rotation therapy provides highly conformal dose distributions.

Discussion and conclusions

Multiobjective inverse planning provides information such as the possibilities (in terms of tumour coverage and protection of OARs and NT) of the dose distributions which are attainable for a range of number of beams. The treatment planner is *not* required to specify unknown information such as importance factors, number of beams and their orientation. Only the prescription dose and the range of the number of beams to be considered has to be specified.

A zero critical dose value is specified for the OARs, thus the aim is to minimize the dose in the OARs by considering dose values even below those which would be considered as clinically acceptable. The main constraint is a sufficiently large coverage of dose for the PTV. Other constraints can be included if desired, such as DVH-based constraints for the OARs. This will increase the number of clinical acceptable solutions.

Methods have been proposed [2, 4] to obtain *optimal* importance factors that provide solution with DVHs very similar to specified *ideal* DVHs. These methods assume some *a priori* knowledge such as the optimal DVHs for the surrounding normal tissue NT and OARs. The problem is that the required dose distribution in external beam radiotherapy cannot always be obtained. This is due to physical limitations and to the existence of trade-offs between the various conflicting optimization objectives. Without multiobjective optimization the treatment planner does not know if the solution found is truly the best possible solution.

As we have a combinatorial problem we do not expect to obtain a global Pareto optimal set. The comparison of the Pareto front obtained by *L-BFGS* and *NSGA-IIc* and the beam orientations found for the test case suggests that the non-dominated set which is obtained is close to the global optimal Pareto front even if only a very small subset of beam orientations can be considered.

The algorithm is not limited to any specific set of objective functions so long as sufficiently efficient algorithms exist that can be used for the optimization of the intensity profile. This is a problem common for all other methods. The role of the MOEA algorithm is to provide the mechanism to select optimal orientations such that a representative set of non-dominated solutions is obtained that is not limited to convex functions and that is not influenced by local minima. We have chosen dose variance based objectives in order to use the very efficient *L-BFGS* algorithm.

For the prostate case we applied soft constraints only on the PTV coverage and the used an off-line filter to select solutions that if possible satisfy all constraints.

We analyzed the dependence of the solutions that passed the filter on the basis of the number of beams. The results showed that the only significant benefit of using larger number of beams is the reduction of the dose in the NT. Especially for this case only a few beam directions are significant to protect the OARs. Only for the NT a larger number of beams are necessary but for more than 6 beams the gain in NT protection seems to be less significant.

For the head and neck case we observed that the PTV coverage and the OAR constraints require a large number of beams. For this case the main problem is that the spinal cord sets a very strict problem. The larger the number of beams is the better the spinal cord can be protected. Using 9 or even 12 beam directions we have to accept a low PTV coverage in order to protect the spinal cord.

The aim of the method presented in this paper is the decoupling of the optimization from the decision making process. Conventional inverse planning assumes some *a priori* knowledge of the possibilities and the treatment planner express this in terms of importance factors, constraints and number of beams. The reality is that planners sometimes devote hours to trying to obtain an acceptable solution by a trial and error change of the importance factors and the constraints.

Our approach is to use some *not too strict* constraints in order to obtain a representative set of all possible solutions. The planner only has to study the trade-off, the limits and the DVHs of the solutions. This can be studied as a function of the number of beams from which the planner can choose a solution that requires a small as possible number of beams.

The multiobjective approach requires the calculation of a large number of solutions to obtain a representative set of solutions. Further studies have to consider the possibilities of reducing this set to the minimum possible size. Meyer et al [27] recently applied influence diagrams for prostate IMRT plan selection using a non-dominated set to analyze the dependence on the importance factors. Romeijn et al [28] studied the possible equivalence of various treatment plan evaluation criteria when used to formulate a multi-criteria IMRT fluence optimization problem.

Additional tools should be provided that help to analyze this set efficiently. Can a method be used that does an automatic analysis of this set and provides the final solution? The answer is not known as long as there is no systematic comparison of the automatic selected solution using the representative set of efficient solutions. Even if so as Webb [29] states:

"Would it be a good thing if a genuinely automated customization technique could be created? At first sight the answer might seem to be affirmative. However, after a while, the skills of human judgment would cease to propagate. Planners might even forget what controls the goodness of outcome. The planning task could become a turnkey. It could become dangerous. Hence I would argue that complete automation is not a desirable objective."

Webb [29] also states that:

"Automation of treatment planning should not remove the human from key decision making".

Multiobjective optimization is considering exactly this important aspect.

Michael Lahanas, PhD
Department of Medical Physics & Engineering
Strahlenklinik
Klinikum Offenbach
Starkenburgring 66
63069 Offenbach am Main, Germany
e-mail: mlahanas@gmx.de

References

1. Peres CA and Brady LW. *Principles and practice of radiotherapy*. 3rd edition. Philadelphia: Lippincott-Raven; 1998.
2. Xing L, Li JG, Donaldson S, Le QT et al. Optimization of importance factors in inverse planning. *Phys Med Biol* 1999; 44: 2525-36.
3. Yu Y. Multiobjective decision theory for computational optimization in radiation therapy. *Med Phys* 1997; 24: 1445-54.
4. Wu X and Zhu Y. An optimization method for importance factors and beam weights based on genetic algorithms for radiotherapy treatment planning. *Phys Med Biol* 2001; 46; 1085-99.
5. Pareto V. Manual of Political Economy. A.M. Kelley, New York, 1971. Translation of Manuale di economia politica, 1906.
6. Miettinen KM. *Nonlinear Multiobjective Optimisation*. Boston: Kluwer Academic Publisher; 1999.
7. Lahanas M, Karouzakis K, Giannouli S, Mould RF et al. Inverse planning in brachytherapy: radium to high dose rate 192 Iridium afterloading. *Nowotwory J Oncol* 2004; 54: 195-218.
8. Stein J, Mohan R, Wang X-H et al. Number and orientation of beams in intensity modulated radiation treatments. *Med Phys* 1997; 24: 149-60.
9. Haas OCL. *Optimization and control systems modelling in radiotherapy treatment planning*, PhD Thesis, Coventry University, 1997.
10. Haas OCL. *Radiotherapy Treatment Planning: New System Approaches, Advances in Industrial Control Monograph*. London: Springer Verlag, 1999.
11. Haas OCL, Burnham KJ, Mills JA. Optimization of beam orientation in radiotherapy using planar geometry. *Phys Med Biol* 1998; 43: 2179-93.
12. Horn J, Nafpliotis N. Multiobjective optimization using the niched Pareto genetic algorithm, IlliGAL Report No.93005, Illinois Genetic Algorithms Laboratory. University of Illinois at Urbana-Champaign, 1993.
13. Schreibmann E, Lahanas M, Uricchio R et al. A geometry based optimization algorithm for conformal external beam radiotherapy. *Phys Med Biol* 2003; 48: 1825–41.
14. Baltas D, Kolotas C, Geramani K, Mould R F et al. A conformal Index (COIN) to evaluate implant quality and dose specification in brachytherapy. *Int J Rad Oncol Biol Phys* 1998: 40; 515-24.
15. Cotrutz C, Lahanas M, Kappas C et al. A multiobjective gradient based dose optimization algorithm for conformal radiotherapy. *Phys Med Biol* 2001; 46: 2161-75.
16. Liu DC, Nocedal J. On the limited memory BFGS method for large scale optimization. *Mathematical Programming* 1989; 45: 503-28.
17. Lahanas M, Schreibmann E, Baltas D. Multiobjective inverse planning for intensity modulated radiotherapy with constraint-free gradient-based optimization algorithms. *Phys Med Biol* 2003; 48: 2843-71.
18. Lahanas M, Schreibmann E, Milickovic N et al. Intensity modulated beam radiation therapy dose optimization with multiobjective evolutionary algorithms,In: Proceedings of the second international conference EMO 2003, Faro, Portugal, April 2003 edited by C. M. Fonseca et al, Springer, pp. 648-661. Lecture Notes in Computer Science. Vol. 2632.
19. Pugachev A, Li JG, Boyer AL, Hancock SL et al. Role of beam orientation optimization in intensity modulated radiation therapy. *Int J Radiation Oncology Biol Phys* 2001; 50: 551-60.
20. Pugachev A and Xing L. Computer-assisted selection of coplanar beam orientations in intensity-modulated radiation therapy. *Phys Med Biol* 2001; 46: 2467–76.

21. Pugachev A, Xing L. Incorporating prior knowledge into beam orientation optimization in IMRT. *Int J Radiat Oncol Biol Phys* 2002; 54: 1565-74.
22. Rowbottom CG, Khoo VS, Webb S. Simultaneous optimization of beam orientations and beam weights in conformal radiotherapy. *Med Phys* 2001; 28: 1696-1702.
23. Rowbottom CG, Webb S, Oldham M. Improvements in prostate radiotherapy from the customization of beam directions *Med Phys* 1998; 25: 1171-9.
24. Meedt G, Alber M, Nüsslin F. Non-coplanar beam direction optimization for intensity-modulated radiotherapy. *Phys Med Biol* 2003; 48: 2999–3019.
25. Schreibmann E, Lahanas M, Xing L, Baltas D. Multiobjective evolutionary optimization of number of beams, their orientations and weights for IMRT. *Phys Med Biol* 2004; 49: 747–70.
26. Deb K and Goel T. Controlled elitist non-dominated sorting genetic algorithms for better convergence, In E. Zitzler et al. (eds.): Evolutionary Multi-criterion Optimization (EMO 2001), *First International Conference, EMO 2001*, Zurich, Switzerland, March 7-9 2001, Proceedings. Lecture Notes in Computer Science Vol. 1993, Springer, 67-81.
27. Meyer J, Phillips MH, Cho PS ET AL. Application of influence diagrams to prostate intensity-modulated radiation therapy plan selection. *Phys Med Biol* 2004; 49: 1637–53.
28. Romeijn HE, Dempsey JF, Li JG, A unifying framework for multi-criteria fluence map optimization models, *Phys Med Biol* 2004; 49: 1991–2013.
29. Webb S. *Intensity-modulated radiation therapy*. IOP Bristol and Philadelphia: Series in Medical Science; 2001.

Paper received and accepted: 7 June 2004