

3D dose computation algorithms

T Knöös^{1,2}

¹Department of Haematology, Oncology and Radiation Physics, Skåne University Hospital, SE-205 02 Malmö, Sweden

²Department of Medical Radiation Physics, Clinical Sciences, Lund University, SE-221 85 Lund, Sweden

Email: tommy.knoos@mod.lu.se

Abstract. The calculation of absorbed dose within patients during external photon beam radiotherapy is reviewed. This includes the modelling of the radiation source i.e. in most cases a linear accelerator (beam modelling) and examples of dose calculation algorithms applied within the patient i.e. the dose engine. For the first part - the beam modelling, the different sources in the treatment head as target, filters and collimators etc are discussed as well as their importance for the photon and electron fluence reaching the patient. The consequences of removing the flattening filter, which several vendors now have made commercially available, is also shown. The pros and cons regarding different dose engines ability to consider density changes within the patient will be covered (*type a* and *b* models). Engines covered are, for example, pencil-beam models, collapsed cone superposition/-convolution models and combinations of these, as well as a glimpse on Monte Carlo methods for radiotherapy. The different models' ability to calculate dose to medium (tissue) and or water is. Finally, the role of commissioning data especially measurements in today's model based dose calculation is presented.

1. Introduction

Clinically used dose planning systems have for many years used calculation algorithms for X-rays and γ -beams, which make use of empirically, determined inhomogeneity corrections. These corrections have been applied to either measured dose distributions using e.g. film or simple analytical models describing the distribution in homogeneous water. Modelling of the dose distribution has not relied on basic photon and electron interaction (basic principles). The methods in use today are often model based algorithms where one separates the modelling of the radiation source from the in-patient dose calculation, the so-called dose engine [1].

2. Historical algorithms

A short description of older algorithms seems to be in place and we divide them depending on their ability to model scattered radiation. These models can be summarised as correction models applied to the dose distribution in water.

2.1. Algorithms without Scatter Modelling

In this group of algorithms, e.g. effective attenuation, isodose shift, tissue air ratios (TAR), effective source skin distance (SSD) [2] the only influence of the patient geometry is through a scaling of the radiological depth for the calculation of absorbed dose. More or less sophisticated versions exist - approximations of homogeneous media, regional corrections or pixel by pixel corrections. These



methods only account for inhomogeneities with respect to their density along the fan line. A better method was introduced with the algorithm by Batho where the position of the inhomogeneity also is considered [3]

These methods were all, in principle, developed prior to the use of CT. The resulting dose distributions from these models can in some clinical cases differ up to 20 % from measurements, especially in low density media irradiated with narrow ($5 \times 5 \text{ cm}^2$) high energy photons beams [4]. Deviations of the same magnitude between measurements and calculations were found for a tangential beam geometry using the effective attenuation method. Of these methods, the Batho method estimates the absorbed dose with the highest accuracy. Lulu and Bjärngard have extended the Batho algorithm to account also for the lateral extension of inhomogeneities [5, 6].

2.2. Algorithms with Scatter Modelling

The next group of methods all use the radiological depth to determine the primary dose component but also accounts for scatter produced in the irradiated volume.

The Equivalent Tissue Air Ratio, ETAR algorithm relies on a calculation of an effective homogeneous density which is assigned to each point using weighting factors applied to the surrounding points (voxels) [7]. Geometric dimensions (e.g. depth, field size) are scaled from unit density to the effective density using O'Connor's theorem [8, 9]. The weighting factors used are calculated using inelastic scattering cross sections for water, i.e. basic principles are applied. The agreement with measurements are better for this method than for the simple algorithms. An accuracy of $\pm 3 \%$ can be expected in most cases [7], however, in some situations, the deviation for the ETAR method can increase, for example, tangential beams, where deviations up to 5 % can be found [10].

An extension of the ETAR method using TAR's calculated by Monte Carlo have been proposed [10]. The new TAR's are more accurate especially for small fields where they fall to zero for zero field size. Therefore, the absence of lateral electronic equilibrium for narrow beams in low density media is modelled better. The final equation used is a superposition integral similar to the methods discussed by Mackie *et al* [11] and Mohan *et al* [12].

In general, none of the methods discussed deals with the lack of electronic equilibrium for narrow beams of high energy X-rays in low density media. All these models assume that the energy transferred to electrons is absorbed locally i.e. the collision kerma is equal to the absorbed dose.

3. Modern dose calculation algorithms

3.1. Multi source models

The limited accuracy of the methods described above has initiated much work on methods based on physical principles. The approach used mainly today is fluence or energy fluence multi source models. The latter is describing the radiation source, usually the treatment head of the accelerator down to different detailed levels. For example, the target generating x-rays from the impinging electron which is considered the primary radiation beam, scattered radiation from primary collimator, beam flattening filter (if present), monitor chamber, multi leaf collimator and conventional jaw system. A common simplification is to reduce all scatter sources to one single usually placed at the position of the flattening filter. This position is chosen since the flattening filter accounts for the majority of the scattered radiation from the treatment head.

Once the model of the radiation source is set up, one calculates the output from it, as an energy fluence distribution. In the case of using Monte Carlo this can be a phase space of the energy fluence i.e. a description of type of particles, their position, direction, and energy. The energy fluence distribution is modulated by field shaping parts of the treatment head. The sources lateral distribution and relative magnitude is usually adjusted based on measurements of conventional measurements e.g. lateral profiles, depth doses, output factors in air and water.

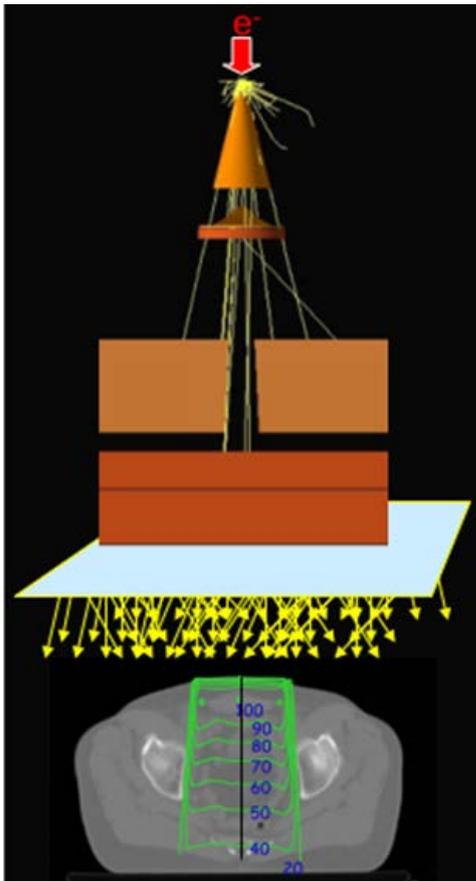


Figure 1. A schematic describing the separation of the modelling into a multi-source model for the treatment head and an in-patient dose calculation. The plane below the jaw system represents the phase space of out-coming particles. The patient is represented with a single transversal cross-section and isodose lines are added for illustration.

3.2. In-patient or dose engine

Dose calculation is based on the previous energy fluence impinging on and into the patient. The patient is represented by a 3D voxel matrix where each voxel is representing the density (relative physical or electron density depending on system). For some models, also a medium may be given for each voxel.

3.2.1. Point kernels

The energy fluence is ray traced thru the patient matrix and attenuated considering the density of each voxel and when appropriate also the medium of the voxel. For each voxel, the TERMA (Total Energy Released in Mass) is calculated as the product of the energy fluence Ψ and the mass energy absorption coefficient μ/ρ . The following equation gives the TERMA T differential in energy E at an arbitrary point in the 3D room given by the vector \mathbf{r} .

$$T_E(\mathbf{r}) = \frac{\mu_E}{\rho} \cdot \Psi_E(\mathbf{r})$$

From the TERMA distribution one gets the absorbed dose by a convolution with a point kernel describing the transport of energy by primary electrons release (collision kerma) and scattered photons. (The differential in energy has been excluded for clarity, the integral has to be performed over all energies).

$$D(\mathbf{r}) = \int T(\mathbf{r}') \cdot P(\mathbf{r} - \mathbf{r}') d\mathbf{r}'$$

The differential in energy has been excluded for clarity, the integral has to be performed over all energies. It is, however, quite common to perform the integral over energy before the convolution, thus only one TERMA matrix and one single point kernel is required.

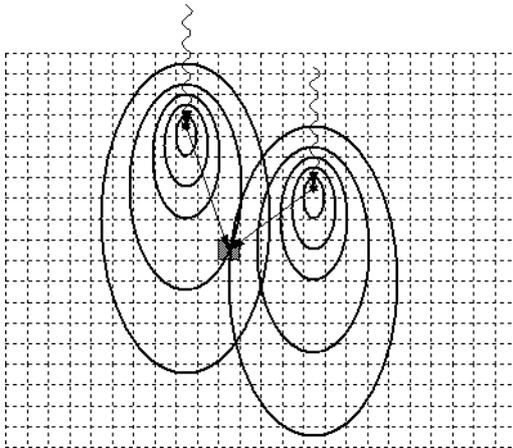


Figure 2. A matrix of TERMA (dotted lines) is convolved with a point kernel, shown as isolevels (solid lines and broken lines respectively) representing the pattern of absorbed dose around an interaction point. The absorbed dose to the shaded voxel (dose point) is determined as the TERMA at the interaction point times the level of the kernel at the dose point. This is repeated for all interaction sites (only two is shown for clarity) to get the total absorbed dose to a single voxel.

This time-consuming integral can be solved using Fourier transforms where the convolution integral is replaced by the inverse Fourier transform of the product between the transforms of T and P . This approach is fundamental in the work by Boyer who used Fast Fourier Transform (FFT) to solve the convolution of dose distribution kernels with photon fluence distributions to give the final dose in 3D [13, 14]. However, if the kernel function P varies with position due to changes in density, an analytical superposition integral has to be solved instead.

The point dose kernels used, which includes electrons released in the first interaction as well as single and multiple scattered photons, are generally calculated using Monte Carlo simulations but analytical methods are also applicable.

3.2.2. Convolution in inhomogeneous medium

Inhomogeneous media can be included in the convolution integral either using a large number of dose distribution kernels, one per density, or a single kernel combined with the following *scaling* theorems:

"In a medium of a given composition exposed to a uniform flux of primary radiation (such as X-rays or neutrons) the flux of secondary radiation is also uniform and independent on the density of the medium as well as of the density variations from point to point." [15]

"In a given irradiation system the ratio of scattered photons to primary photons remains unchanged when the density of the irradiated material is changed if all the linear dimensions of the system are altered in inverse proportions to the change in density." [8, 9]

Density in this context should be interpreted as interaction sites per unit volume, e.g. electron density for incoherent scattering of photons. It is also assumed that the amount of secondary radiation produced at a point is proportional to the local number of interaction sites.

Mackie *et al* used superposition kernels calculated for several densities with interpolation and scaling between them [11]. A similar approach using kernels for only a *single* density combined with the scaling theorems has been applied [12, 16-18]. They all use ray tracing of scatter between interaction and absorption point and employ the scaling theorem for inhomogeneities. Transportation of energy by electrons is included in the kernels, therefore, electronic disequilibrium will be modelled.

3.2.3. Pencil beam kernels

A method to increase the calculation speed to make it clinical feasible was the introduction of pencil beam kernels where the dose calculation problem has decreased by going from 3D to 2D convolution. By integrating the point kernel along the depth direction, we get a 2D description of the dose distribution.

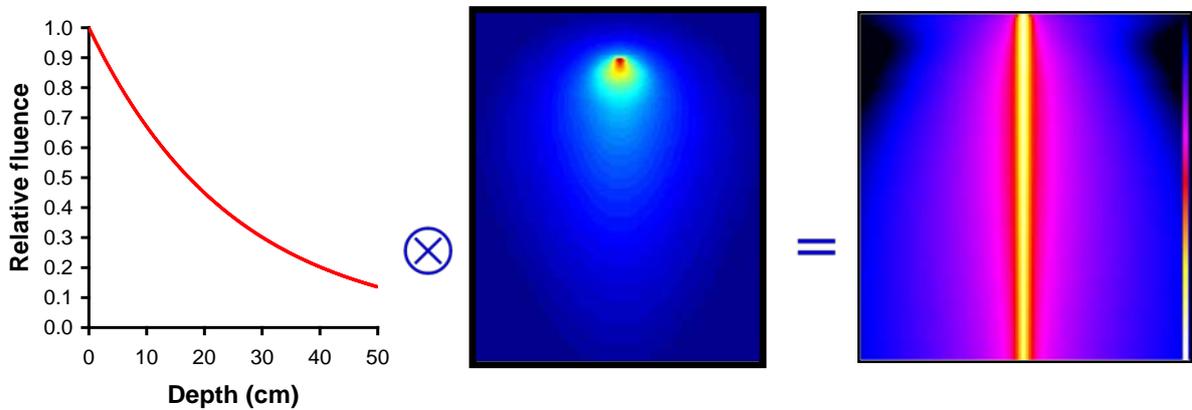


Figure 3. Illustration of the principle of pencil beams showing creation of pencil beams by convolving/integrating TERMA along a ray with the point kernel describing the energy transport and absorption. Left panel - TERMA along the central ray of a very narrow photon beam, middle panel - point kernel and right panel - the resulting pencil beam.

Pencil beams have been created in various ways for example from measurements by radial differentiation of relative dose on central axis from broad beam dosimetric quantities [19]. Another method was the differentiation of radial beam data [20]. Monte Carlo methods have also been used to calculate pencil beam kernels in water [16, 21, 22]. Usually the pencil beams are parametrized and the actual values of these are found by a fitting process to measured data [22].

Applying the pencil beam convolution is described in principle by the following equation:

$$D(x, y, z) = \iint \Psi(x', y') \cdot P(x - x', y - y', z) dx dy$$

Here the energy fluence distribution Ψ at a certain specified plane is convolved with the pencil beam kernel P . In heterogeneous media, the pencil beam kernel is scaled along the propagation direction by replacing the z with the radiological depth z_{radiol} considering the density of the voxels along the ray. No scaling is performed, in the initial implementations, perpendicular to the propagation direction. Thus, the pencil beam convolution is a so-called *type a* algorithm [23]. There is, however, today one pencil beam model where a lateral scaling of the kernel has been added [24]. The commercial implementation is the analytical anisotropic algorithm (AAA) [25, 26]. The performance of pencil beams algorithms have been reported in several reports [27-32].

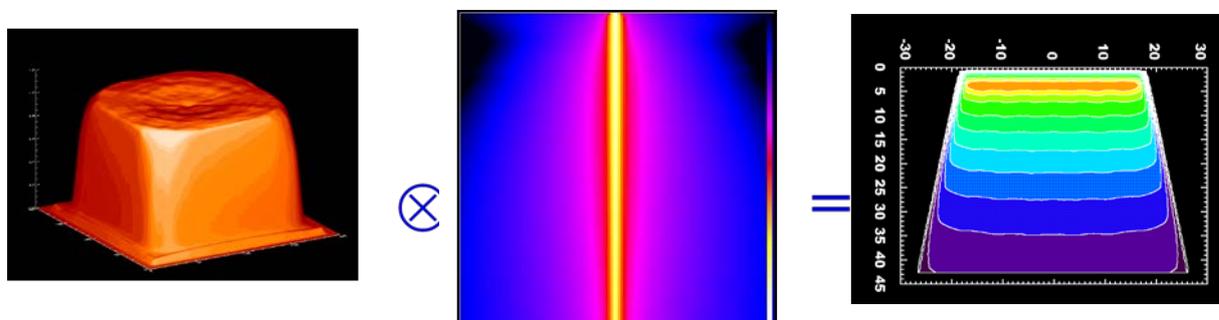


Figure 4. Illustration of the 2D energy fluence distribution (left) convolved with the pencil beam (middle) to get the dose distribution represented by a transverse plane (right).

3.2.4. Collapsed cone convolution

Another approximation to increase speed (which maybe is not a problem today) as well as managing inhomogeneities is to discretize the point kernel. The kernel is discretized in about 100 directions and by letting each direction representing a cone where all energy is collapsed to the center of the cone [33, 34].

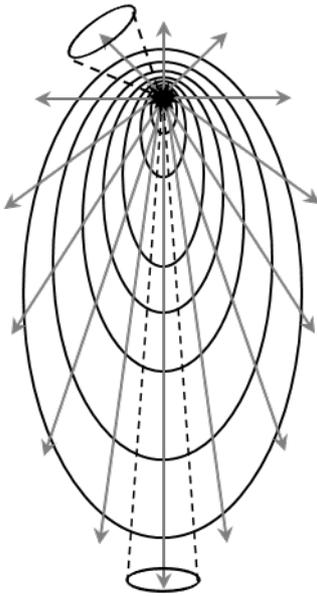


Figure 5. The point kernel (illustrated by the isolevels) is discretised along the arrows which each represent a cone where energy is transported. All energy in the cone is collapsed to the central ray.

When collapsing the energy in the cone into the central ray the inverse square dependence is removed. The distribution of cones is not isotropic, instead a higher density is used in the forward direction because the majority of energy in high energy beams is transported in this direction. A few cones take care of the backscatter. Density of cones should be high enough where significant amount of energy is deposited such that each voxel is passed by one or more collapsed cones.

The convolution integral is commutative thus changing the order of the functions does not change the result

$$D(\mathbf{r}) = \int T(\mathbf{r}') \cdot P(\mathbf{r} - \mathbf{r}') d\mathbf{r} \text{ is equal to } D(\mathbf{r}) = \int P(\mathbf{r}') \cdot T(\mathbf{r} - \mathbf{r}') d\mathbf{r}$$

This property can be interpreted as shown in the following figure:

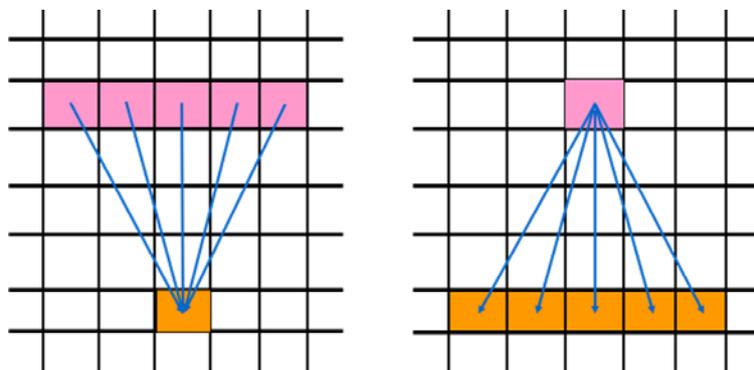


Figure 6. Illustration of the commutative property of the convolution integral. The left panel shows the "dose deposition view" and the right panel the "interaction view" [4].

For the dose deposition view, energy or TERMA from the surrounding voxels is summed up to the deposition voxel. When all contributions are summed up, the total dose in the voxel is determined. In the interaction view, the TERMA in a voxel is spread out to the surrounding voxels and all contributions from interaction points have to be covered. In principle, the first approach can be used to calculate the absorbed dose in a single point/voxel only. The implementations based on this approach utilize the possibility to have a dose grid with different spacing to speed up calculations. For example, in areas with small gradients in TERMA or patient density a coarse grid can be used and in high gradients a finer one.

Another approach used to efficiently perform the convolution is to align the point kernels such that the cones in the same direction overlaps. This makes it possible to ray trace through the TERMA matrix along each cone direction and recursively pick up energy from each traversed voxel, transport, attenuate and then deposit energy along the axis of the cone [33].

3.3. Linear Boltzmann transport equation

A more direct way of performing in-patient dose calculations is to start with the linear transport equation by Boltzmann (LBTE). For many years one have used Monte Carlo (MC) techniques to statistically solve the problem [35-38]. MC methods use stochastic distributions describing the physics to simulate the transport of individual particles (photons and charge particles) through the density and medium matrix describing the patient geometry. The limitation, at least up today, is the large number of particles is required to get a sufficient high uncertainty of the deposited dose in each voxel of interest. MC can be seen as a microscopic solution of the LBTE.

Recently analytical solutions have been applied to LBTE for radiotherapy planning problems in a macroscopic approach i.e. instead of individual particles we create, follow, and attenuate energy fluencies in a matrix describing the patient. An exact analytical solution is not possible, mainly due to its complexity, thus a numerical have to be used, i.e. discrete ordinate method or grid based solvers [39-42]. Evaluation of a commercial implementation have been presented by e.g. Fogliata *et al* [43, 44] and Hoffman *et al* [45].

In short, the grid based method can be described as; the properties of the radiation source and the absorbing medium (patient) are described, the multi-source model (e.g. by a phase space) is used to get the energy fluence to the patient and then ray traced through the irradiated volume (observe this is very similar to the point kernels methods discussed above). In the next step the discrete ordinates is applied to solve the LBT equation giving photon energy fluence and the electron sources in each voxel. Finally, the electrons are transported to obtain electron fluence distribution in each voxel and by applying mass stopping power the absorbed dose is determined. The energy of photons and electrons are discretized, 25 and 49 levels, retrospectively. Transport directions is discretized based on the energy involved between 32 and 512 directions (These data are from the Acuros implementation [46]).

Interesting to comment is that both solutions to the LBTE show inherent errors, for MC we have stochastic uncertainties. For discrete ordinate solvers, the concept of discretization may produce uncertainties. As always this can be improved by prolonging the computation by increasing the number of particles in the MC case and for the solver using a finer discretization.

4. Dose to medium or dose to water

For many years, there have been a discussion on how to report absorbed dose during radiotherapy procedures [47]. No consensus really exists if the dose should be reported as absorbed in water or in the actual medium. With new algorithms as MC and grid based LBTE solvers getting more common this problem is growing. Both these approaches can report either.

The issue can be divided into two parts, a) if the energy/TERMA during the ray trace considers the physical properties of the transversed voxels or all media is considered to be water, and b) the other step is how the deposition of energy is accomplished. Introducing the formalism $D_{\text{transport, deposit}}$ where transport and deposit are the medium used respectively. They can either take the value water (w) or medium (m).

Considering the kernel approaches, pencil beam and point kernels, are all determined in water, analytical, MC or from experiments. Looking in detail to pencil beam algorithms, scaling according to the transversed density (radiological depth scaling) assumes all media is water and no corrections are performed during dose deposition. Thus pencil beam models transports and deposit dose in water, $D_{w,w}$.

For the point kernel models available, no consistent handling exists. In principle, the TERMA or equivalent can be ray-traced considering the medias physical properties or not. The kernels are determined in water but when dose deposition occurs a correction using mass stopping power ratio between medium and water may be applied. No single conclusion on what is reported can be given, however, we probably have one of the following situations: $D_{w,w}$, $D_{m,m}$ or $D_{m,w}$.

MC models have the possibility to transport in water or medium and the same is valid for dose deposition. Most common is to transport in medium and deposit in medium. If the user want dose to water, a mass stopping power ratio between water and medium is applied to each voxel [48]. Thus we have $D_{m,m}$ or $D_{m,w}$. When applying the mass stopping power ratio one usually have to use a macroscopic value for the energy deposit in the voxel since the electron energy fluence differentiated in energy in the voxel is not known. Alternatively, this could be done microscopically if done for each single energy deposition.

The discrete ordinate solvers for the LBTE determines the electron energy fluence differentiated in energy for each voxel and either the mass stopping power for the medium or for water is used to determine the absorbed dose. All transport is accomplished in the medium thus we have $D_{m,m}$ or $D_{m,w}$. The choice is user selectable in the only existing commercial implementation [49, 50].

5. Summary

Today we have very accurate tools available for estimating the absorbed dose within the patient during radiotherapy [51-57]. We still have, however, several problems to solve. The models discussed here are in principle for static patients without movements i.e. intrafractional and/or interfractional movements. The latter can probably be solved by daily imaging and adapting the today's plan. This can be accomplished by e.g. a library of plans or on-line re-optimisation and calculation. For intrafractional movements, we have used PTV, ITV etc. and margin recipes to assure that the target get the correct dose. This will in many cases lead to an over-irradiation of healthy tissues surrounding the tumour. Techniques with synchronisation of the patient with the treatment delivery e.g. gating and tracking is close to be clinical routine at many departments.

Another problem that has been noticed when introducing more accurate algorithms is a tendency of optimizers to put in too much dose in the vicinity of the tumour to assure an adequate dose in the whole tumour. This is especially noted when working with the PTV concept and tumours in low density regions such as the lung. In these cases, one should probably use a combination of the *type a* and *b* algorithms where optimisation is performed using *type a* and then followed by a recalculation with an appropriate *type b* model. More physically one can describe this as an optimisation based on energy fluence (pencil beam model) followed by dose calculation with full lateral electron modelling.

6. References

- [1] Ahnesjö A and Aspradakis M A 1999 *Phys. Med. Biol.* **44** R99-155
- [2] ICRU 1976 **24** Bethesda, Maryland, USA
- [3] Batho H F 1964 *J. Canadian Ass. Radiologists* **15** 79-83
- [4] Mackie T R *et al* 1985 *Med. Phys.* **12** 188-96
- [5] Lulu B A and Bjärngard B E 1982 *Med. Phys.* **9** 372-7
- [6] Lulu B A and Bjärngard B E 1982 *Med. Phys.* **9** 907-9
- [7] Sontag M R and Cunningham J R 1978 *Radiology* **129** 787-94
- [8] O'Connor J E 1957 *Phys. Med. Biol.* **1** 352-69
- [9] O'Connor J E 1984 *Med. Phys.* **11** 678-80
- [10] Wong J W and Henkelman R M 1983 *Med. Phys.* **10** 199-208
- [11] Mackie T R *et al* 1985 *Med. Phys.* **12** 327-32
- [12] Mohan R *et al* 1986 *Med. Phys.* **13** 64-73

- [13] Boyer A 1984 *Med. Phys.* **11** 552-4
- [14] Boyer A and Mok E 1985 *Med. Phys.* **12** 169-77
- [15] Fano U 1954 *Radiat. Res.* **1** 237-40
- [16] Ahnesjö A *et al* 1987 *Acta. Oncol.* **26** 49-56
- [17] Knöös T and Nilsson M 1987 *Proc. 9th Int. Conf. on the Use of Comp. in Rad. Ther.*
- [18] Nilsson M and Knöös T 1992 *Phys. Med. Biol.* **37** 69-83
- [19] Ceberg C P *et al* 1996 *Med. Phys.* **23** 505-11
- [20] Storchi P and Woudstra E 1996 *Phys. Med. Biol.* **41** 637-656
- [21] Mohan R and Chui C S 1987 *Med. Phys.* **14** 70-7
- [22] Ahnesjö A *et al* 1992 *Med. Phys.* **19** 263-73
- [23] Knöös T *et al* 2006 *Phys. Med. Biol.* **51** 5785-807
- [24] Ulmer W *et al* 2005 *Med. Biol.* **50** 1767-1790
- [25] Van Esch A *et al* 2006 *Med. Phys.* **33** 4130-48
- [26] Tillikainen L *et al* 2008 *Phys. Med. Biol.* **53** 3821-39
- [27] Knöös T *et al* 1994 *Phys. Med. Biol.* **39** 1609-28
- [28] Knöös T *et al* 1995 *Phys. Med. Biol.* **40** 1411-20
- [29] Knöös T *et al* 1995 *Radiother. Oncol.* **37** 74-80
- [30] Weber L *et al* 1996 *Med. Phys.* **23** 307-16
- [31] Hurkmans C *et al* 1996 *Phys. Med. Biol.* **41** 1277-90
- [32] Hansson H *et al* 1999 *Phys. Med. Biol.* **44** N201-7
- [33] Ahnesjö A 1989 *Med. Phys.* **16** 577-92
- [34] Miften M *et al* 2000 *Phys. Med. Biol.* **45** 817-33
- [35] Nelson W R *et al* 1985 *Stanford Linear Accelerator Report SLAC-265*
- [36] Fippel M 2004 *Med. Phys.* **31** 1235-42
- [37] Rogers D W 2006 *Phys. Med. Biol.* **51** R287-301
- [38] Chetty I J *et al* 2007 *Med. Phys.* **34** 4818-53
- [39] Wareing T A *et al* 2001 *Nucl. Sci. Eng.* **138** 256-68
- [40] Gifford K A *et al* 2006 *Phys. Med. Biol.* **51** 2253-65
- [41] Vassiliev O N *et al* 2008 *Int. J. Radiat. Oncol.* **72** 220-7
- [42] Vassiliev O N *et al* 2010 *Phys. Med. Biol.* **55** 581-98
- [43] Fogliata A *et al* 2011 *Radiat. Oncol.* **6** 82
- [44] Fogliata A *et al* 2011 *Med. Phys.* **38** 6228-37
- [45] Hoffmann L *et al* 2012 *Acta Oncologica* **51** 376-85
- [46] Failla G A *et al* 2015 *Varian white paper RAD 10156 10/10 350*
- [47] Ma C M and Li J 2011 *Phys. Med. Biol.* **56** 3073-89
- [48] Siebers J V *et al* 2000 *Phys. Med. Biol.* **45** 983-95
- [49] Mampuya W A *et al* 2016 *J. Appl. Clin. Med. Phys.* **8** 6338
- [50] Rana S *et al* 2014 *Int. J. Cancer Ther. Oncol.* **2** 020421
- [51] Mather M L *et al* 2003 *Ultrasonics* **41** 551-9
- [52] Baldock C *et al* 2010 *Phys. Med. Biol.* **55** R1-63
- [53] Vial P *et al* 2006 *Phys. Med. Biol.* **51** 5517-38
- [54] Hurley C *et al* 2006 *Nucl. Instrum. Meth. A* **565** 801-11
- [55] Baldock C 2009 *J. Phys.: Conf. Ser.* **164** 01200
- [56] Hill B *et al* 2005 *Med. Phys.* **32** 1589-97
- [57] Hill R *et al* 2009 *Med. Phys.* **36** 3971-81