ORIGINAL ARTICLE

Taylor&Francis healthsciences

Determination and Clinical Verification of Dose-Response Parameters for Esophageal Stricture from Head and Neck Radiotherapy

Panayiotis Mavroidis, Göran Laurell, Thomas Kraepelien, Jan-Olof Fernberg, Bengt K. Lind and Anders Brahme

From the Departments of Medical Radiation Physics, Karolinska Institutet and Stockholm University (P. Mavroidis, B.K. Lind, A. Brahme), Otolaryngology, Head and Neck Surgery, Karolinska Hospital (G. Laurell), Hospital Physics, Radiumhemmet, Karolinska Hospital (T. Kraepelien) and Oncology, Radiumhemmet, Karolinska Hospital (J.-O. Fernberg), Stockholm, Sweden and the Department of Medical Physics, Larisa University Hospital, (P. Mavroidis), Larisa, Greece

Correspondence to: Panayiotis Mavroidis, Department of Medical Radiation Physics, Karolinska Institutet and Stockholm University, PO Box 260, S-171 76 Stockholm, Sweden. Tel: +46 8 5177 2225. Fax: +46 8 343525. E-mail: panayiotis.mavroidis@radfys.ki.se

Acta Oncologica Vol. 42, No. 8, pp. 865–881, 2003

The purpose of this work is to determine the parameters and evaluate the predictive strength of the relative seriality model. This is accomplished by associating the calculated complication rates with the clinical follow-up records. The study is based on 82 patients who received radiation treatment for head and neck cancer. For each patient the 3D dose distribution delivered to the esophagus and the clinical treatment outcome were available. Clinical symptoms and radiological findings were used to assess the manifestation of radiation-induced esophageal strictures. These data were introduced into a maximum likelihood fitting to calculate the best estimates of the parameters used by the relative seriality model ($D_{50} = 68.4$ Gy, $\gamma = 6.55$, $s = 0.22$). The uncertainties of these parameters were also calculated and their individual influence on the dose-response curve was demonstrated. The best estimate of the parameters was applied to 58 patients of the study material and their esophageal stricture induction probabilities were calculated to illustrate the clinical utilization of the calculated parameters. The calculation of the biological effective dose (BED) appeared to be significantly sensitive to the applied fractionation correction for complex treatment plans. The relative seriality model was proved suitable in reproducing the treatment outcome pattern of the patient material studied (probability of finding a worse fit = 61.0%, the area under the ROC curve = 0.84 and χ^2 test = 0.95). The analysis was carried out for the upper 5 cm of the esophagus (proximal esophagus) where all the strictures are formed. Radiation-induced strictures were found to have a strong volume dependence (low relative seriality). The uncertainties of the parameters appear to have a significant supporting role on the estimated dose-response curve.

Received 18 December 2002
Accepted 22 May 2003

Impairment of the digestive function is commonly seen in patients with head and neck cancer. The wide range of swallowing problems induced by therapeutic interventions is generally considered as a drawback that usually has to be accepted. Injury to the gastrointestinal tract is a common complication after radiotherapy for pelvic and abdominal malignant tumors. The incidence of gastrointestinal irradiation injury depends on many factors but the knowledge on normal tissue radiosensitivity is weak for most sites (1, 2). The esophagus at the upper end of the gastrointestinal tract is reported to be fragile and radiosensitive (3). Acute radiation esophagitis is commonly seen in connection to radiotherapy for operable breast cancer, inoperable non-small cell carcinoma of the lung and esophageal carcinoma (4–6).

Information about the dose-response relations of tumors and normal tissues has been reported by many scientists (7–15). Newer and more accurate clinical data that are being published at an increasing rate make evaluation of biological treatment planning more reliable (1). However, to introduce information about normal tissue radiation reactions in the clinical routine, it is necessary to quantify it through clinical trials (1, 16–19). The radiobiological normal tissue response, which correlates the dose delivered to the tissue with its radiosensitivity, is quantified through the use of radiobiological models. These models describe the dose-response relationships of tumors and normal tissues by using the three-dimensional dose distribution of the applied treatment plans and radiobiological parameters that describe the radiobiological characteristics of these

 ISSN 0284-186X
DOI: 10.1080/02841860310012833

organs. In this sense, they are an important tool in the calculation of normal tissue complication probabilities (NTCP) and consequently in the evaluation and comparison of treatment plans (20–23). Clinical implementation of NTCP modeling can be achieved gradually, assuming that the dose-response curves have accurately reached the point of describing the radiation induction of a specific end-point for a specific organ.

To perform radiotherapy optimization using radiobiological data in the case of head and neck cancer radiotherapy, relevant radiobiological parameters must be estimated from clinical material. The parameters used by the different radiobiological models are generally derived from patient data sets, so that they have some clinical validity. These kinds of studies are carried out retrospectively when the clinical outcome of the treatment is known. In the present analysis, the parameters characterizing the dose-response relation of radiation-induced esophageal strictures following radiotherapy for head and neck cancer are derived. Patient material consisting of 82 cases having individual esophageal responses and doses to esophagus was analyzed using the relative seriality model. The best estimates of the model parameters were determined by a maximum likelihood fitting. The parameter confidence intervals around the best estimates were subsequently calculated.

The estimated radiobiological parameters refer usually to a certain dose per fraction. For this reason, the dose distributions delivered to the patients have to be converted to an equivalent fractionation regime that delivers that dose per fraction. The linear-quadratic model is applied often to account for those fractionation effects. The application of the proper fractionation correction may have a significant effect on the calculated dose distribution. For instance, when different treatment configurations are applied in different fractions then the fractionation correction has to be applied separately to each one of them before calculating the combined dose distribution.

The patient material was selected without any bias concerning the radiotherapy treatment techniques used. Inter-patient and intra-patient esophageal radiosensitivity was assumed to be the same throughout the study. Radiobiological parameters describing the radiosensitivity of almost every organ can be found in the literature (1, 17–19). However, only few studies illustrate their clinical utilization and provide means to the readers for examining whether certain parameters are applicable to their local organ delineation methods and treatment techniques (24). This kind of study can be sensitive to the treatment methodology applied, especially for normal tissues where the dose distribution delivered to them can vary significantly between different irradiation techniques. For this reason, many details have to be provided in order to help other radiotherapy centers verify whether the derived parameters are compatible with their clinical methodology and patient characteristics. Furthermore, a clinical application of the derived parameters is demonstrated. The probability of stricture induction following radiation therapy is calculated for a subgroup of the patients. The calculated probability is then compared with the known esophageal radiation response of this group to illustrate their close association. However, introduction of radiobiological modeling into the clinical practice requires analysis of further independent studies, using the same endpoint definition.

MATERIAL AND METHODS

Patient selection

The patient group consisted of 26 patients with radiation-induced esophageal complications and 56 having no symptoms (control group). These 82 patients were selected from the overall population of 925 cases treated with radiotherapy for head and neck cancer between 1992 and 2001 at Radiumhemmet, Karolinska Hospital, Stockholm. The treatment modality was decided by a team of radiation oncologists, head and neck surgeons and medical physicists prior therapy. Almost a third of the patients were treated with (pre-operative) radiotherapy followed by surgery. A similar proportion of patients underwent first surgical treatment followed by radiotherapy and the rest of the patients received radiotherapy as the only therapeutic modality. The present study was carried out having radiation-induced esophageal stricture as the clinical endpoint. The assessment of the radiation injury was done retrospectively through diagnosis of the clinical symptoms, together with radiological findings. The symptoms underlying the clinical diagnosis were swallowing problems and weight loss. The patients' medical records were reviewed up to 7 years after radiation treatment for data collection. The interval between the end of radiotherapy and the diagnosis of stricture was 1–40 months (median 7 months, mean 9 months). Out of 26 patients, 21 formed their strictures within the first 10 months after radiotherapy. Data about smoking history of the patients before radiotherapy were too poor to allow a valid statistical analysis. The clinical characteristics of this patient group, which also refer to age, gender and tumor sites are summarized in Table 1. The analysis of this work is based on an accurate delineation of the organ during the treatment planning and on an accurate dose delivery to the esophagus according to the treatment plans.

Treatment techniques and dose planning

All patients were set up in supine position and immobilized by a patient dedicated mask. This position was held throughout the Computer Tomography (CT) acquisition, treatment simulation and complete treatment regimen. Treatment planning was performed on a three-dimensional

Table 1

Description of the clinical, treatment and follow-up characteristics of the patient population. Distribution of the patients according to tumor site and primary tumor stage is presented (the numbers in the parenthesis represent number of patients without deglutition disorder whereas the rest of the numbers correspond to patients with proximal esophageal stricture; three patients were not classified)

Clinical characteristic	All	Complications	Complication-free
Patients	82	26	56
Males	48 (58.5%)	14	34
Females	34 (41.5%)	12	22
Age			
Range	14–92 years	14–81 years	40–92 years
Mean	63.2 years	59.8 years	64.5 years
Median	62.5 years	60.5 years	64 years
Irradiation technique			
Anterior angled fields	11 (13.4%)	0	11
Unilateral fields	24 (29.3%)	3	21
Bilateral fields	47 (57.3%)	23	24
Follow-up recording	–	Conventional swallow radiograph Endoscopy	Questionnaire survey
Other clinical symptoms	–	Impaired swallowing function Weight loss	Symptom-free

Time	Follow-up	Stricture formation
Range	13–86 months	1–40 months
Median	62.0 months	7 months

Tumor site	T1	T2	T3	T4	Total
Oral cavity	1 (2)	3 (7)	4	1 (4)	22
Glottis	(8)	(5)	(1)	(1)	15
Epipharynx	2	–	–	1	3
Oropharynx	(5)	(4)	2 (2)	2 (3)	18
Hypopharynx	1	–	(1)	–	2
Larynx	(3)	3	3	–	9
Miscellaneous	2 (4)	(2)	(2)	–	10

(3D) conformal treatment planning system (TMS, Helax MDS-Nordion). The planning was based on approximately 30 contiguously acquired CT slices with thickness of 0.5–1.0 cm. Corrections for tissue inhomogeneities were also performed in the treatment planning process.

All patients in the study were treated with 6 MV photons. The treatments were performed through isocentric conformal techniques where fields were individually shaped by means of custom-made blocks and multileaf collimators (MLC). The blocking of the spinal cord was always performed with cerrobend block mounted on the tray-holder to achieve optimal penumbra. In some cases, the volume and the weight of the cerrobend block were reduced by using a multileaf collimator as a backup collimation.

Clinical Target Volume (CTV) consisted of the primary gross tumor (GT) and the locally involved lymph nodes (LN) (25). A margin of 1–2 cm in all directions from known or suspected tumor was included to form the Planning Target Volume (PTV). Cervical and supraclavicular lymph nodes were included in all treatments except for early glottic cancers (T1). In bilateral treatments the posterior neck nodes were treated to 46 Gy but otherwise a dose of 64 Gy was delivered to the whole neck and supraclavicular nodes. Treatments were lateralized in cases of lateral primary tumor without neck nodes or with only ipsilateral spread. In hypopharyngeal and laryngeal cancers, treatment was always bilateral.

Patients with small tumors in frontal localization (glottic larynx T1) were treated in a standardized way with two oblique, tilted rectangular fields without shielding (Fig. 1). Patients with laterally located targets were planned with one frontal and one oblique posterior field (cf Fig. 2). The treatment technique for a bilateral internal target volume (ITV), which included a primary tumor, bilateral lymph nodes and supraclavicular lymph nodes, consisted of two opposed coplanar or non-coplanar lateral conformal fields (Fig. 3). In some of the cases where it was difficult to cover the supraclavicular region with the prescribed dose, the region was treated with anterior and posterior fields. The junction between the lateral fields and the anterior-posterior fields was at least 1–2 cm from the gross tumor volume (GTV). To improve the homogeneity of the dose distribution in a volume ranging from mandibula to jugulum, low weighted beam segments were added within the large fields.

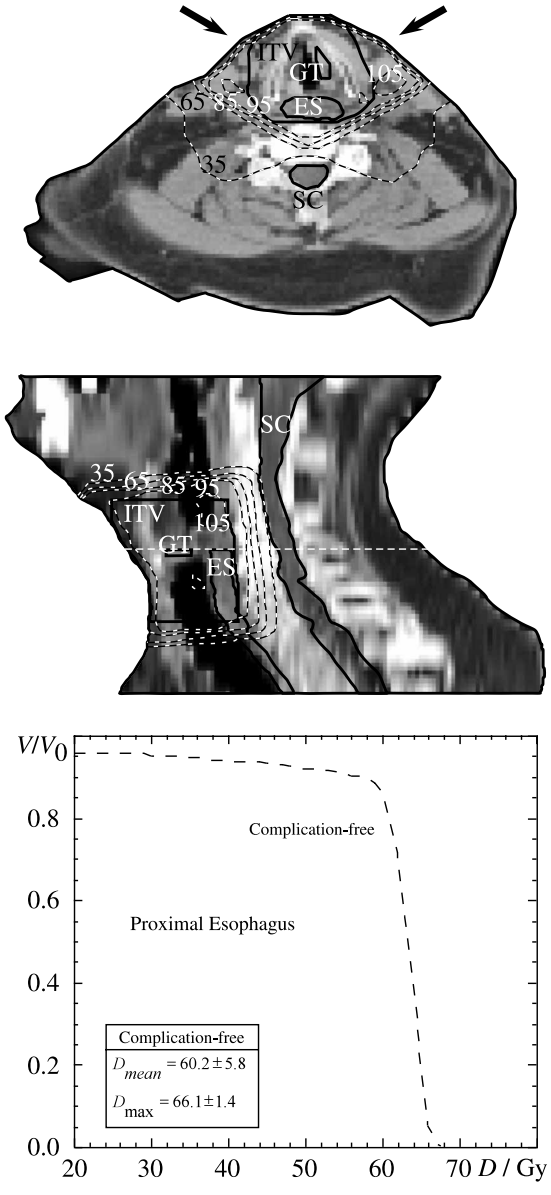

Fig. 1. In this treatment technique, two angled anterior fields are applied. This is a treatment configuration that is applied to clinical cases of small tumors without lymph node involvement. In the transverse slice the anatomical structures of the gross tumor (GT), the esophagus (ES) and the spinal cord (SC) that are involved in the clinical case are shown. The transverse image refers to the slice where the proximal esophagus first appears. This can also be seen from the sagittal view, which shows that the proximal esophagus lies in the high dose region. The dose delivered to the upper 5 cm of the esophagus is presented by the mean cumulative dose volume histogram. In this diagram, the notation 'Complication-free' refers to the fraction of those patients treated with this technique who did not have any complications. For this technique none of the patients showed any complication after radiotherapy.

The following guidelines, specified by the radiotherapists, were followed in treatment planning:

- Doses of preferably 65–70 Gy should be given to the GTV and similar doses (60–70 Gy) should be delivered to the lymph nodes apart from the posterior nodes, which should receive 46 Gy.
- The dose to the spinal cord must not be higher than 50 Gy.

The dose was delivered by a fractionation schedule of 2 Gy per fraction, five fractions per week. These treatment techniques were used routinely at Radiumhemmet, Karolinska Hospital during the period of the study. DVHs were calculated for the upper part of the esophagus, which is the most relevant to the clinical findings. In the cases of bilateral ITV, the treatment volume was treated to a dose of 46 Gy in 23 fractions over 4.5 weeks. The volume anterior to the spinal cord was then boosted to 64 Gy in nine supplementary fractions. The absorbed dose to the spinal cord did not exceed its (routinely used) tolerance dose of 50 Gy. For patients with lateral ITV or small frontal ITV (larynx, stage I), the complete ITV was covered with 64 Gy and the dose to the spinal cord still did not exceed 50 Gy.

In vivo dosimetry was performed according to the ordinary routine on all fields in every fraction. The results of the dose measurements from this study were within accepted limits ($\approx 2.9\%$ deviation between the calculated and the delivered dose per field). A better agreement between the planned and the delivered dose could be achieved through the methods of Lind (26), Lind et al. (27) and Löf (28). In cases where eye lens were located close to the treated volume, the absorbed dose to these organs at risk was determined by means of Thermo Luminescence Dosimetry (TLD, Harshaw QS 5500).

Follow-up

All patients had a nursing visit every two weeks during and after radiotherapy at the ward for head and neck cancer, department of Otolaryngology. After the final treatment the patients were seen every 1–3 months by the surgeon. Clinical symptoms that lead to suspicion of a stricture were impaired swallowing function and weight loss. These symptoms were either persistent after radiotherapy or with a late onset. Thirteen patients needed nutritional support but were able to swallow solid food of reduced pulp. Another eight patients were unable to swallow solid food but could still drink. Finally, five patients presented severe dysphagia and were unable even to drink water without aspiration. Esophageal strictures were graded in increasing order of severity from I to III according to the following definitions: I. Dysphagia to solid food but not to pulp/semisolids. Moderate stricture (moderate fibrosis) could

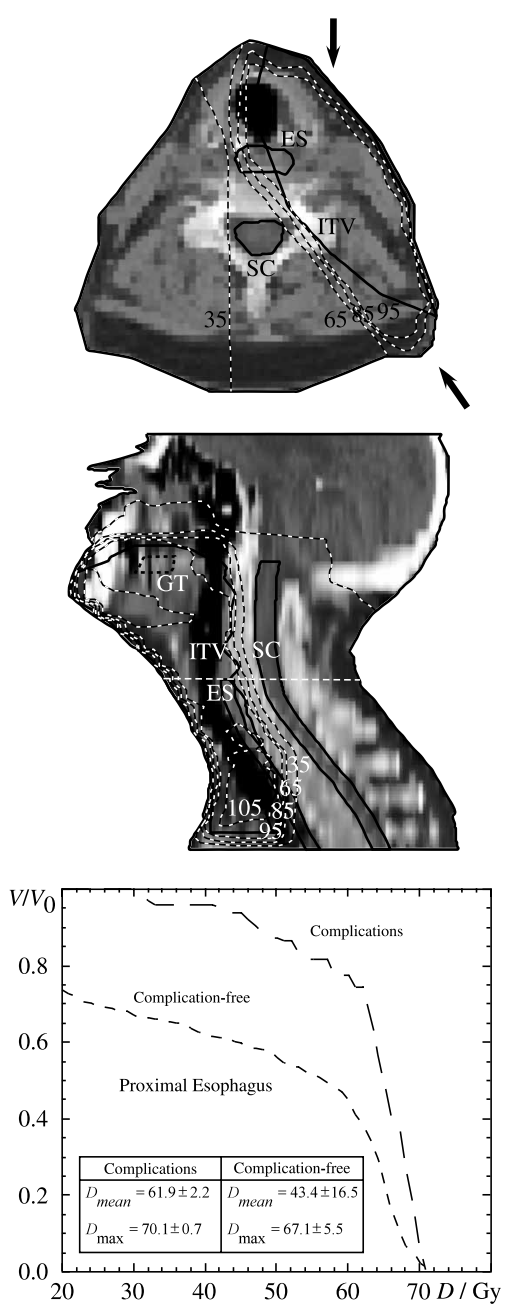

Fig. 2. In this treatment technique, the applied fields irradiate only one side of the patient's neck. This is a treatment configuration applied to clinical cases where the tumor and the lymph node involvement lie laterally. In the transverse slice the internal target volume (ITV) is shown together with the anatomical structures of the normal tissues involved. Although in the sagittal view it is shown that the proximal esophagus lie in the high dose region, the transverse view shows that the dose fall-off takes place inside the esophagus, meaning that parts of it receive low dose. This can also be observed in the lower diagram, which illustrates the mean cumulative dose volume histogram of the upper 5 cm of the esophagus. The term 'Complications' corresponds to the fraction of the patients treated with this technique, who had radiation-induced stricture as complication.

initially be passed by a rigid esophagoscope of 7 × 10 mm and dilated; II. Dysphagia to solid food but not to liquids. In the severe stricture (severe fibrosis) of the esophageal inlet it was not possible to pass the stricture using the smallest endoscope without dilatation; and III. Total obliteration where there were no visible communication between hypopharynx and esophagus in the endoscopy (29–32). In this study, the diagnosis of any of these grades of dysphagia has been considered as complication.

Conventional swallow x-ray using barium or water-based contrast material was performed as the initial examination in 23 of the complication cases. Radiographs illustrating esophageal strictures having different degrees of severity are presented in Fig. 4. Endoscopy under general anesthesia was applied to the injured patients to diagnose the degree of the stricture. The radiotherapeutic data for the patients with esophageal stricture were compared with 56 head and neck patients without dysphagia after radiotherapy. A survey that was carried out by mailing a questionnaire to all surviving patients receiving radiotherapy for a head and neck tumor in 1994 at the department of Oncology, Radiumhemmet was used for selecting all the 22 patients without any swallowing problems. The remaining 34 consecutive cases were all complication-free patients that were treated in 2000–2001 for head and neck tumors at the same radiotherapy center.

The relative seriality model

The complications observed in normal tissues following the therapeutic use of radiation have been described in terms of inactivation of functional subunits (FSU). The volume effect describes how the tolerance dose increases with decreasing partial volume of the normal tissue being irradiated. Calculating the probability of causing injury to normal tissues is much more complicated than for tumors since it is dependent on the internal structure and organization of the functional subunits of the irradiated organ. To determine how damage to functional subunits leads to the expression of complications it is important to understand how organs are functionally structured in parallel and serial subunits.

In this work, the concept of the volume effect is treated by the model of relative seriality (7, 19–21). For a heterogeneous dose distribution, the response of normal tissues is given by the expression:

$$P_{\mathrm{I}}(\vec{D}) = \left[1 - \prod_{i=1}^{M} (1 - \exp(-e^{e\gamma - (D_i/D_{50})\cdot(e\gamma - \ln(\ln(2)))})^{s})^{\Delta\gamma_i}\right]^{1/s} \quad [1]$$

where $P_{\mathrm{I}}(\vec{D})$ (I denotes injury) is the probability of inducing a certain injury to an organ that is irradiated with a dose distribution $\vec{D}$. D_{50} is the dose which gives a response probability of 50% and γ is the maximum normalized value of the dose-response gradient. Both D_{50} and γ depend on the initial number of functional subunits for healthy tissues.

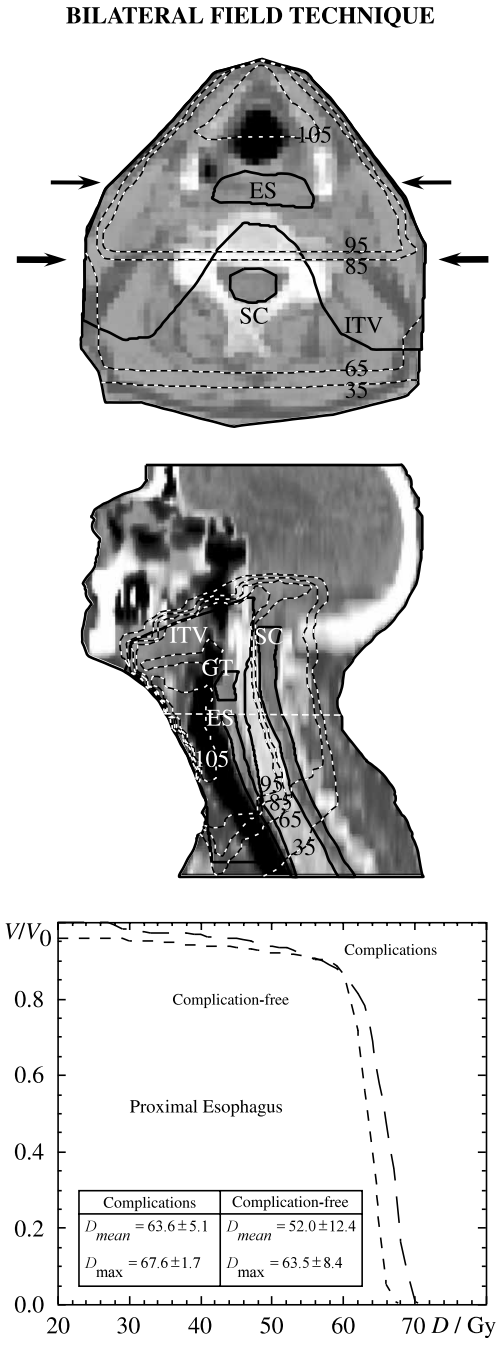

Fig. 3. In this treatment technique, two different treatment configurations are combined. Both of these dose plans involve lateral fields. The first configuration is applied for 23 fractions (2 Gy per fraction) though the second one follows for nine fractions. The main difference between the two configurations is that a protective block is introduced in the way of the beams in the second plan to protect the spinal cord. However, the dose distribution in the esophagus is approximately the same for both of the plans. From the transverse and the sagittal views it is shown that the proximal esophagus lies in the high dose region. This is also depicted in the diagram showing the mean cumulative dose volume histogram of the upper 5 cm of the esophagus.

Δv_i ($=\Delta V_i/V_{\text{ref}}$) is the fractional subvolume of an organ that is irradiated compared to the volume of the reference length for which the values of D_{50} and γ were calculated. s is the relative seriality parameter that characterizes the internal organization of the organ and M is the total number of subvolumes in the organ.

The radiobiological model for cell kill that was used in this work for describing the dose-response relation of esophagus regarding the endpoint of radiation-induced stricture is the linear-quadratic-Poisson model (7, 8, 12, 19, 20, 23), which also accounts for the fractionation scheme applied. Each esophageal dose distribution was corrected to a 2 Gy per fraction using the linear-quadratic model (33). Thus, each dose step in the dose volume histograms was corrected separately. The α/β value assumed in the linear-quadratic model correction was 3 Gy (34). However, a sensitivity test was performed and the calculation was repeated using α/β values up to 8 Gy even though there is no evidence in the literature indicating a deviation from the value of 3 Gy. The radiation sensitivity was assumed to be homogeneous throughout the esophageal volume.

Different approaches in treating the concept of reference volume for the normal tissues have been reported. Usually, the whole volume of the healthy organ is considered as reference volume, because the volume of an organ is related to the functional needs of the individual human being. In this study, the upper 5 cm of the proximal esophagus has been used as the reference length on the grounds that the esophageal strictures are formed in this region. Moreover, the techniques applied deliver high dose only to this part of the organ, with the rest of it receiving significantly reduced dose. The use of shorter esophageal length would become critical in terms of migration of epithelia cells from adjacent (marginal) mucosa as the dose at the margins would not be trivial in such a case.

Maximum likelihood fitting of the response model to clinical data

The values of the radiobiological parameters used by the model and their confidence intervals were determined through a fitting of the theoretical predictions for complications (applying the dosimetric information of each patient to the response model) to the clinical follow-up results. The group of patients with complications is considered to be those patients who showed esophageal stricture of any degree of severity. The point of patient selection is very essential in studies where dose-response relations are derived. In most of these studies, the sampling method that is followed considers a number of consecutive patients whose treatment and follow-up information is known. In this approach, however, it is possible to use inaccurate proportions of the different irradiation techniques applied in the clinic and inaccurate incidence rates, and also to

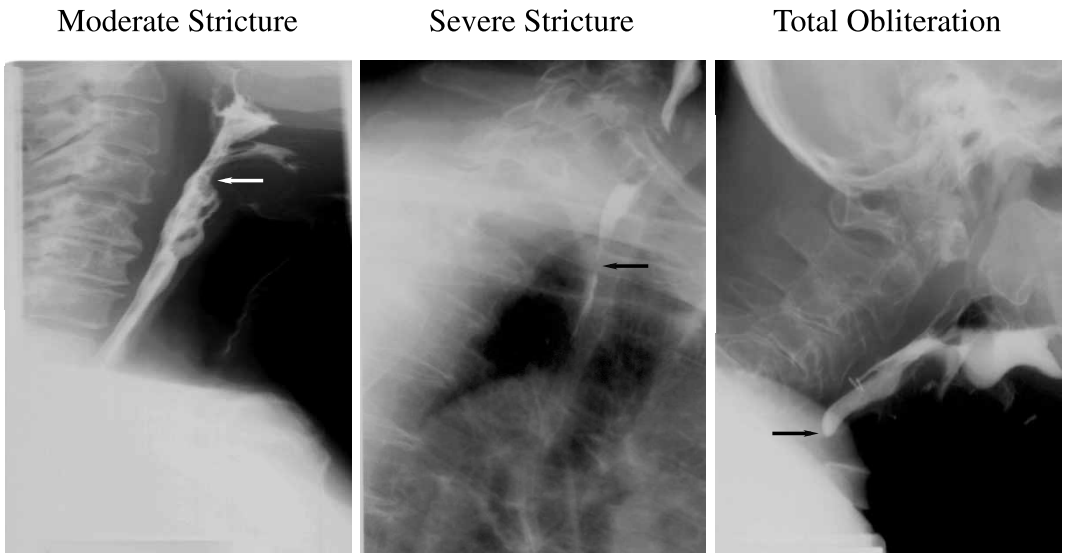

Fig. 4. Radiographs showing three representative cases of patients with radiation-induced strictures. They follow the classification order of increasing severity: moderate stricture, severe stricture and total obliteration. The radiographs were produced using barium as the contrast media. The arrows in the graphs point out the parts of the esophagus where the strictures have been formed.

overlook factors like sex, age and surgery, which may have a larger or smaller impact on the radiobiological modeling and parameter derivation processes. When many such factors are involved, a case-control approach seems to be a more appropriate one since the characteristics of the overall patient population are known and are taken into account accordingly. This holds under the assumption that the patient sample has to be large enough to cover a considerable period of time and represent the overall population treated in the clinic. In this study, the two approaches have been combined appropriately. In order to have a realistic incidence rate, the material was rescaled to conform to the clinical incidence rate ($\sim 7\%$, based on the records of the last 12 years). This was done by increasing the weight of the complication-free group so that the final proportion of the two groups was of the order of the clinical rate. This course of action was accurate for three reasons: first, the dosimetric characteristics of the patients without complications were fairly similar for each treatment technique; second, the proportions of the three treatment techniques in the study group are the same as those in the clinic; and third, the estimation of the dose-response curve depends on the requirement that the sample subgroups of the patients with and without complications are individually representative of the large patient population that has been treated in the clinic and from which the study samples have been derived. For the fitting of the parameters the maximum likelihood method was used, which determines the best estimates of the parameters by maximizing the likelihood to reproduce the given pattern of observations (18, 19, 35, 36).

The fitting calculations were performed through the use of a minimization package, MINOS (37). For the present calculations the parameter space was restricted to positive values. To find the global maximum of the logarithm of the likelihood function avoiding eventual local maxima, the calculations were performed by changing both the initial values (starting points) and the allowed range of the parameters. The goodness of the fit was estimated by applying three independent statistical methods.

Evaluation of the goodness of fit of the dose-response parameters

According to the method described by Jackson *et al*. (36) and Eadie et al. (38), given D_{50}, γ and s from the maximization result, the average of the log-likelihood function and its variance are calculated assuming a Gaussian distribution for the log-likelihood function. The expected mean value and standard deviation are then compared to the optimum value of the log-likelihood function found by the maximization process. Subsequently, the probability of achieving a worse fit regarding the estimated pattern of complications is calculated. A large probability (close to 1) would mean that a very good agreement between the two distributions (predictions, clinical results) has been reached and the fit can be considered as optimal.

The Pearson's χ^2 test is a statistic that characterizes the dispersion of the observed frequencies from the expected frequencies. The numerator of the χ^2 formula is a measure of the spread of the observations whereas the denominator is a good measure of the expected spread. In this work, the goodness of fit is assessed by comparing the observed and

expected complication results in a table where the patient population has been divided in subgroups. Each patient subgroup covers a different dose range of the dose-response curve. By applying the χ^2 test to the different pairs of results (observed against expected) the probability that a random sample of data points drawn from the assumed probability distribution would yield a value of χ^2 as large or larger than the observed value in a given experiment with v degrees of freedom is calculated. If the probability is reasonably large (close to 1), then the predicted distribution correctly describes the spread of the data points. If the probability is small, either the predicted distribution is not a good estimate of the parent distribution or the data sample is not representative of the parent distribution.

Another approach to test the fitting was to perform a validation based on the analysis of receiver operating characteristic (ROC) curves (39, 40), which are plots of the true positive rate against the false positive rate for different possible cut-offs of a diagnostic or predictive test. The individual probabilities of observing an esophageal stricture were calculated using the dose distribution delivered to each patient and the model parameters derived from the maximum likelihood fitting. The patients were ordered by their calculated probabilities and a series of trial cut-offs, P_{cut} was examined. By comparing the number of observed complications above P_{cut} with the number of predicted complications, the *true positive ratio* (TPR, the ratio of observed positive incidences above P_{cut} to the total number of positive incidences) and the *false positive ratio* (FPR, the ratio of observed incidence-free patients above P_{cut} to the total number of incidence-free patients) were determined. TPR and FPR were plotted against each other in the form of a ROC curve. The area under the curve measures discrimination, which is the ability of the test to correctly classify those with and without the disease. For perfect classification of the observed against the predicted complication results, the area under the curve is 1. Random assignment of complication results leads to a ROC area of 0.5.

An evaluation of the parameter errors was done by calculating their 68% and 95% confidence intervals. These were obtained by calculating the three-dimensional 68% and 95% joint confidence regions of the parameters D_{50}, γ and s, i.e. the region in the $D_{50}-\gamma-s$ hyper-volume, in which it is estimated that there is 68% and 95% probability of finding the true values of the three parameters respectively. The $D_{50}-\gamma-s$ hyper-volume was found by varying the parameters around their optimum values. In order to study the impact of the parameter uncertainties on the dose-response curve, a bundle of dose-response curves was calculated using parameter values from the calculated confidence intervals. The described procedure represents a first step in quantifying the uncertainty of the dose-response curve, due to the uncertainties in D_{50}, γ and s, imposed by the fitting procedure.

The biologically effective uniform dose, $\bar{\vec{D}}$

The biologically effective uniform dose is the uniform dose that causes exactly the same tumor control or normal tissue complication probability as the real dose distribution (23, 24, 41).

$$\bar{\vec{D}} = D_{50}\frac{e\gamma - \ln(-\ln(P(\vec{D})))}{e\gamma - \ln(\ln(2))}. \quad [2]$$

In this work, this concept is used to find the uniform dose that is as biologically effective as the dose distribution, $\vec{D}$ delivered to each patient in the study population. For each patient the effectiveness of the applied dose distribution is calculated by the relative seriality model and the set of parameters estimated by the maximum likelihood method.

RESULTS

Two representative samples of 26 patients with esophageal strictures (according to the follow-up definitions) and 56 controls were drawn from their respective populations. The average appearance of the symptoms was between 3 to 6 months after the completion of radiotherapy. The mean age of the patients at radiotherapy time was 59.8 ± 15.3 years (median 60.5 years) in the complication group, versus 64.5 ± 12.6 years (median 64.0 years) in the group without complications. Of the patients that developed esophageal strictures seven had been treated with pre-operative radiotherapy, six had received post-operative radiotherapy and 13 had been treated only with radiotherapy. A statistically significant positive association of radiation stricture induction with dose (cut-off at 64 Gy) was found (odd ratio (OR) = 6.62 with 95% confidence interval (CI) 2.20–19.86). The patients of the overall population did not receive chemotherapy in combination with radiotherapy apart from very few exceptional cases.

The mean cumulative DVHs of the proximal esophagus were obtained for the three different treatment techniques used in this study. In the lower diagrams of Figs. 1–3, these mean cumulative DVHs are illustrated for the first 5 cm of the proximal esophagus of the groups of patients with and without complications. Of the patients 11 were treated with anterior angled fields (none showed complications), 24 were irradiated with unilateral fields (three showed complications) and 47 were treated with bilateral fields (23 showed complications). The averaged mean and maximum doses delivered by the different techniques have also been calculated. This information demonstrates the association of the irradiation configuration with the dose delivered to the esophagus and the clinical outcome of the treatment.

Application of the fractionation correction after the summation of the different dose distributions can introduce significant errors in the calculated dose. In Fig. 5, it is

shown that for the spinal cord this effect was significant because the dose from the two field arrangements to this region differed considerably. However, for the esophagus this effect was negligible since the dose distributions produced by the two dose plans were very similar.

The optimum parameter estimates with their 68% and 95% confidence intervals are given in Table 2. These calculations were performed by correcting each step of the DVHs with the linear-quadratic model and assuming an α/β value of 3 Gy. The sensitivity test, which was performed by using α/β values up to 8 Gy in the fitting procedure, provided parameter values very close to the first ones, showing that the present fitting process is not strongly dependent on the α/β value of the fractionation correction. The optimization did not show any dependence on the parameter limits chosen for the calculation. Different incidence rates (up to 12%) were tried to examine the dependence of the process on the incidence rate used but it was proved that the scaling of the patient subgroups with and without complications did not practically change the estimated values of the model parameters. Furthermore, the stability of the estimated values was tested using samples of the patient subgroups with and without complications, showing that the parameters D_{50} and s remained fairly stable around their best estimates while γ varied ± 1 ($\sim 15\%$) around its corresponding value.

Based on the association between length (volume) and esophageal stricture, the analysis was carried out for the upper 5 cm of the proximal esophagus. The best estimates of the model parameters obtained were $D_{50} = 68.4$ Gy, $\gamma = 6.55$ and $s = 0.22$. However, among the patients that developed esophageal strictures there were four cases where the region of esophagus under study received exceptionally low dose. For these cases it may be that eventual bacterial infections (causing mucositis or candida) or simultaneous surgery close to this region combined with radiotherapy sensitize the esophagus assisting the formation of a stricture. Because of these underlying implications and the fact that such factors may introduce some bias in our results, the radiobiological parameters were calculated excluding these patients from the analysis. In the left diagram of Fig. 6, the mean cumulative DVHs representative of the mean dose distribution delivered to the patients is illustrated. The dose-response curves obtained with the best parameter estimates (Table 2) are given in the right diagram of Fig. 6, where the curves are plotted for different fractions of irradiated esophageal volume. The spread among the dose-response curves corresponding to whole and partial esophagus irradiation indicates that the induction of esophageal radiation strictures varies substantially with the irradiated volume. This is supported by the fact that 49% of the patients treated with the bilateral technique showed complications, though only 27% of those treated with the unilateral technique, which partially irradiates the esophagus, developed strictures. In other words, from a functional point of view the esophagus shows a parallel-like architecture for this endpoint, which is mathematically expressed in the relative seriality model by an s value close to zero.

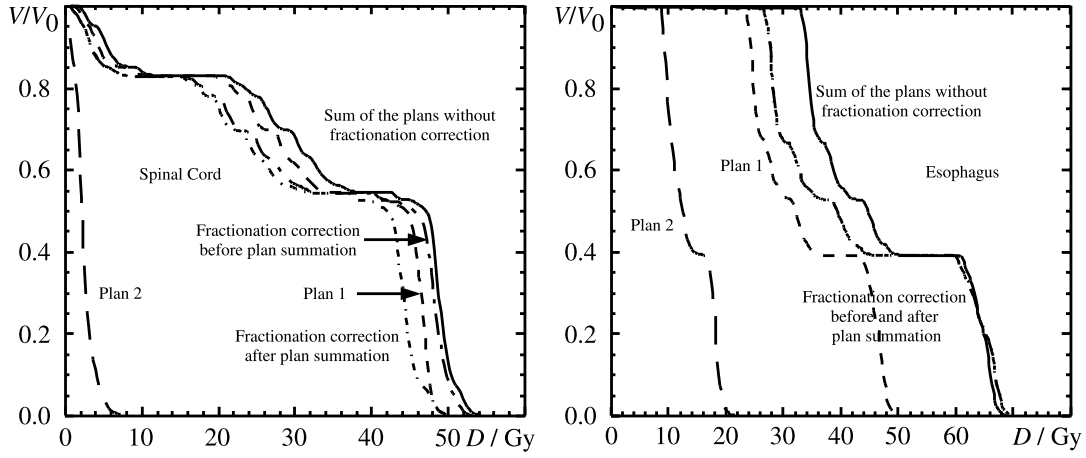

Fig. 5. Illustration of important features of fractionation correction in cases where different treatment configurations are delivered in different fractions, producing a complex treatment plan such as in the case of the bilateral field technique. For the case of the spinal cord, the left diagram shows the cumulative dose volume histograms of the two treatment configurations and their combination: (a) without any fractionation correction of the combined plan; (b) with fractionation correction of the combined plan to 2 Gy per fraction; and (c) with fractionation correction of the separate plans to 2 Gy per fraction before they are combined. The latter is the correct way to make the fractionation correction and from the diagram it can be seen that the method applied can have a significant effect on the calculation of the corrected dose volume histogram. Particularly large discrepancies take place when the different treatment configurations deliver significantly different dose distributions to the site under consideration. In the case of the esophagus (right diagram) it is shown that making the fractionation correction before or after the combination of the individual dose volume histograms does not have a significant effect because the two plans produce very similar dose distributions in the esophagus.

Table 2

Best estimates and confidence intervals of the relative seriality model parameters derived from the patient material of the study

Parameters	Best estimates and confidence intervals		
	$\hat{P}$	68%	95%
D_{50}, dose giving 50% of complication probability (Gy)	68.4	67.0–72.5	66.3–82.1
γ, maximum normalized gradient of the dose-response curve	6.55	3.34–10.22	1.57–11.79
s, relative seriality parameter	0.22	0.088–0.38	0.006–0.55

The 68% and 95% joint confidence regions for the parameters D_{50}, γ and s were obtained by calculating the hyper-volumes corresponding to these probabilities (Fig. 7, left diagram). A bundle of dose-response curves can be obtained by using D_{50} and γ values from their confidence intervals. The range that is defined from these response curves is depicted in the right diagram of Fig. 7 for whole volume irradiation. The choice of representing the confidence interval of the dose-response relation with a bundle of dose-response curves derived from the corresponding $D_{50}-\gamma$ joint confidence region was based on the consideration that any combination of the parameter values within this region can be possible for a patient.

The goodness of fit was determined by the mean (expected value) and standard deviation of the maximum likelihood function used. The expected value of the log-likelihood function obtained from the patient complication probabilities, was calculated using the fitted parameters, and resulted to be -59.39 with a variance of 72.09. The observed value of the log-likelihood function from the fit was -57.68. Assuming a Gaussian distribution of the $\ln L$ (cf Appendix), these results indicate that the probability of finding a worse fit (smaller value of $\ln L$) is 61.0%. Dividing the patient population in 11 subgroups and applying the Pearson's χ^2 test, the value of 0.315 was calculated for χ^2 (Table 3). The probability that stems from this value of χ^2 (for seven degrees of freedom) for having a perfect agreement between the expected and the observed complications results is 0.95. This value indicates that the relative seriality model and the estimated parameters reproduce very well the pattern of the clinically recorded complications. This is also shown by the results of the predicted and observed complication rates (7.37% and 6.98%, respectively). As another measure of the predictive power of the applied radiobiological model, the area under the ROC curve, which measures the ability of the model to classify the patients with and without esophageal strictures, is also calculated. The value of 0.84 of the area under the ROC curve indicates that the model and the derived parameters distinguish quite well the groups of patients with and without complications based on their dosimetric characteristics (Fig. 8, left diagram). This result is also supported by the average values of $\bar{\bar{D}}$ of the patient groups with and without complications, which are 65.2 and 60.6 Gy, respectively.

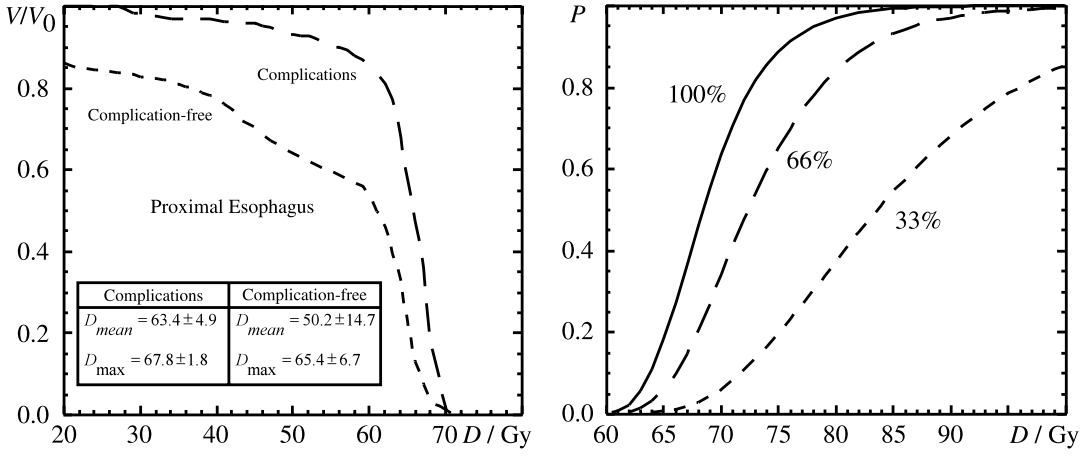

Fig. 6. The mean cumulative dose volume histogram of the whole study population, divided to the complication-free group and the group of patients with radiation-induced strictures for the upper 5 cm of esophagus is shown in the left diagram. Furthermore, a quantitative expression of those histograms is also provided. In the right diagram, the corresponding dose-response curve of the estimated radiobiological parameters is presented for a range of uniform doses. The volume dependence of the parameters is demonstrated by assuming that: (a) the whole organ (100%) receives a uniform dose; (b) two thirds (66%) of the organ receives a certain uniform dose and the remaining part receives 5% of this dose; and (c) one third (33%) of the organ receives a uniform dose and the rest 5% of this dose. It is obvious that by increasing the parts of esophagus being spared, higher doses can be sustained for the same incidence rate of esophageal strictures.

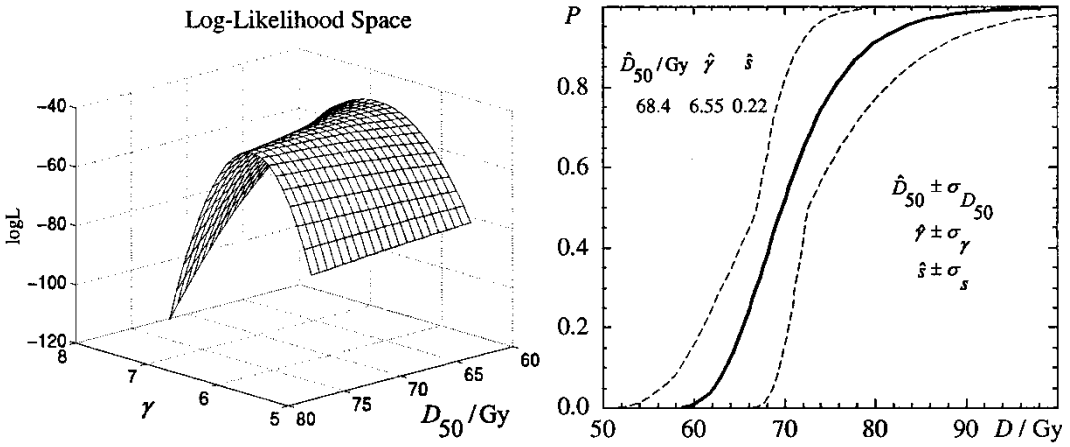

Fig. 7. Left diagram: The hyper-surface shows how the logarithm of the maximum likelihood function changes in the region around the best estimates of the radiobiological parameters. This hyper-surface was calculated using the best estimate of s while varying the rest of the model parameters (D_{50}, γ). The maximum point of the diagram corresponds to the best estimate of the parameters (maximum of the log-likelihood function). From the actual hyper-volume of the log-likelihood function (it is calculated using all the possible combinations of parameter values) the confidence interval of the estimated radiobiological parameters is calculated. Right diagram: The confidence range of the dose-response curve was derived from combinations of D_{50} and γ values (s was fixed to its best estimate) lying within their calculated confidence interval of 68%. This diagram illustrates the uncertainty in the estimated dose-response curve, which stems from the inter-patient radiosensitivities and dosimetric discrepancies between the calculated and the delivered dose. The clinical implementation requires the knowledge of the variability introduced by treatment and patient-specific factors to the prediction of the clinical outcome.

To illustrate the clinical use of the calculated parameters a group of 58 patients was selected from the study population. Using the derived radiobiological parameters for the 5 cm reference length and the relative seriality model, the dose-response curve of esophagus is calculated for a range of uniform doses (Fig. 8, right diagram). Subsequently, the response probability is calculated for every patient using these parameters again and their individual esophageal dose distribution. By applying the concept of biologically effective uniform dose on these probabilities, the corresponding $\bar{\bar{D}}$ values are found for each patient. By definition the points fall exactly on the predicted dose-response curve. To examine whether the theoretical curve reproduces the observed complication rates, the 58 patients confined in the dose ranges (60.0–63.5), (63.5–65.5) and (65.5–67.0) Gy are selected. In the first interval, three patients developed strictures corresponding to an incidence rate of 2.26% (scaled proportion). In the second interval eight patients had complications (11.8% incidence rate) while in the third interval six patients showed strictures (28.6% incidence rate). The average expected response probabilities of the patients in those groups were 2.98%, 13.8% and 27.8% respectively. In this diagram, comparisons between average predicted and true follow-up rates can be made only for dose ranges where a considerable number of patients with and without complications exist (the dose range where the 58 patients are located). In this way, the comparison claims statistical validity. Towards the low and high doses such comparisons become strongly inaccurate because the low number of patients in these regions tends to make the comparison statistically biased. Taking into account the fact that the number of the patients used was small the two sets of probability values are quite close. This is because the selected patients belong to the data set from which the parameters were calculated. Nevertheless, this is a good way to examine if a set of parameters is compatible with a certain treatment technique with the prerequisite that the same endpoint definition is used.

Table 3

Results from the fit of the biological model on the patient data. The goodness of fit was determined for the patient subgroup[1] by different methods (normal error distribution, Pearson's test, ROC analysis)

Statistical method	Association of theoretical and clinical results			
Normal error distribution	Observed log$L = -57.68$	Expected log$L = -59.39$	Variance log$L = 72.09$	$P_{\text{worse-fit}}$ (%) = 61.0
Pearson's χ^2 test	Predicted P_I (%) = 7.37	Observed P_I (%) = 6.98	[2]$\chi^2_v = 0.315$	$P_\chi(\chi^2_v, v) = 0.95$
ROC	0.84			

[1] Four patients were not included in the analysis.
[2] Degrees of freedom = number of patients subgroups (= 11)–number of fitted parameters (= 3)–1.

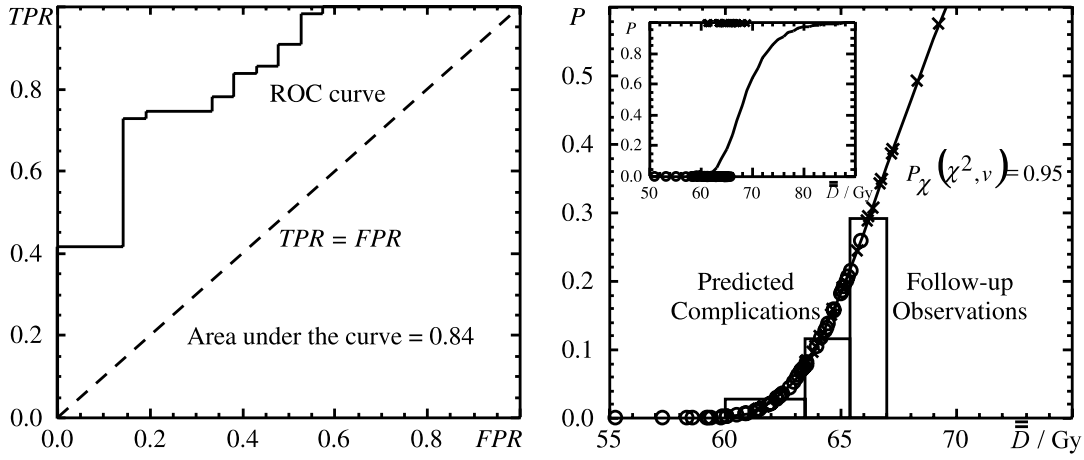

Fig. 8. Left diagram: Using the estimated model parameters to calculate the expected complication probabilities of the patients and ordering them as in the left diagram of the figure, a *receiver operating characteristic* (ROC) curve could be constructed using different cut-off thresholds and calculating the corresponding *true positive ratio* (TPR) and *false positive ratio* (FPR) values. This method evaluates the predictive strength of the model by comparing ratios of the true and false incidences predicted. The area under the curve indicates a good agreement between the expected and the observed complication data. Right diagram: The dose-response curve produced using the estimated radiobiological parameters and the relative seriality model (solid line). The unit of the dose axis is the biologically effective uniform dose, $\bar{\bar{D}}$. By using this dose unit the position of every patient of the study population can be found on the theoretical response curve (crosses for the patients with complications, white circles for the complication-free patients). The small diagram shows the dose-response curve in full range with the patients being placed according to their clinically observed response (0 or 1). Three dose intervals between the $\bar{\bar{D}}$ values of 60 and 67 Gy are observed where 2.26%, 11.8% and 28.6% of the patients had complications. The expected complication rates of the patients in those intervals are 2.98%, 13.8% and 27.8%, respectively, which are fairly close to the clinical observation regarding the small number of patients selected. This was expected since the parameter values used were derived from the same study population. However, it is shown how these parameters should be applied in the clinic and how one could check if some published parameters are suitable for a certain treatment methodology.

DISCUSSION

At the beginning of this work, 42 patients were analyzed, half of which had radiation-induced esophageal strictures (42). A survey using a questionnaire that was mailed to all surviving patients who received radiotherapy during 1994 at the Department of Oncology, Radiumhemmet, Karolinska Hospital, Stockholm for a head and neck tumor, selected all patients without any swallowing problems. The data files showed that irradiation had been administered to 105 patients with head and neck carcinoma during 1994. In January 1996, 65 of these patients were still alive. Of the 65 surviving patients, 51 answered the questionnaire regarding swallowing problems. 22 of them were completely free of swallowing problems and had complete radiotherapeutic dose planning records. Using this material the parameters (D_{50}, γ and s) of the relative seriality model were calculated. The problem at that time was that it was not certain whether the complication-free group was representative of the larger complication-free patient population. By the end of this process it was realized that the number of patients could be increased by including 40 consecutive patients in the study (from 42 to 82). In this way, it would be possible to verify the validity of the first results and make the arguments and results that are derived by the analysis stronger. Conveniently, the new patients supported our first preliminary results. Using now a much larger material, possible sampling inaccuracies could to a large extent be eliminated. This is because the present complication-free group has the same proportions as the large complication-free patient population (1992–2001) between the preoperative, postoperative and without-surgery subgroups as well as between the three treatment techniques. It is believed that these factors may affect the observed radiosensitivity of the patients.

Usually, the calculated radiobiological parameters describing the dose-response relation of an organ refer to a certain uniform dose per fraction. Consequently, the dose delivered to the patients has to be adjusted to this dose per fraction before deriving these parameters. In the first two treatment techniques described, the fractionation correction was straightforward. However, in the bilateral field irradiation, where two different treatment configurations are applied in different fractions, the correction is more complex. In such cases, the different treatment arrangements should be corrected separately for the fractionation effects and then be combined to produce the summed dose distribution. The treatment configurations described in the present study result in a large part of the proximal esophagus being irradiated to a high dose, especially the upper 5 cm. Generally, it is desirable to have many different treatment techniques in the study material because this usually implies that the clinical data cover a larger dose range on the dose-response curve (since the dose distributions may differ significantly) leading to a more accurate

estimate of the curve. Moreover, in this case systematic errors that stem from the applied treatment technique smooth out. The radiobiological parameters derived from such analyses refer usually to a certain dose per fraction (most often 2 Gy). So, a correction for the fractionation effects has to be applied to the dose plans before using their dosimetric information. In this study, the fractionation correction was applied using the linear-quadratic model. Although this model is accurate for large doses it has not been validated for doses lower than 1 Gy. Consequently, the correction may be approximate in this dose region. Moreover, when different treatment configurations are used in different fractions of the same treatment plan, as in the case of the bilateral field technique, the fractionation correction is more complex. The different treatment configurations have to be corrected separately for the fractionation effects before they are summed up to form the final dose distribution. The effect of the appropriate fractionation correction on the calculated dose volume histogram is shown in Fig. 5.

For this clinical endpoint, the actual incidence of complications is typically in the order of $\sim 7\%$ (clinical records). Radiation-induced esophageal strictures have been investigated by only a few groups because of the difficulty in collecting uniform data. The different factors related to such an analysis, such as endpoint definition, follow-up recording, adequate dosimetric data and the different radiobiological models used, make comparisons of data from different sources a complicated task. The lack of data uniformity in radiobiological modeling is partly dependent on the fact that most studies (such as the present one) are based on a retrospective assessment of the treatment outcome. The quality and the amount of available information have thus been the limiting factors in the definition of the endpoint and in its measurement. Quantification of the dose distribution in the esophagus has often been approximate because of the absence of 3D dose planning.

A very interesting point of this study is that the derived relative seriality (s value) is substantially lower than previous estimates (1, 17). There are two main reasons why this deviation occurs. Firstly, the dose delivered to the esophagus in many of the previous studies was not known to the extent that it is now (3D treatment planning systems, etc). Secondly, almost all these studies investigated another part of esophagus than the one that this study is dealing with (proximal esophagus). The difference observed supports the argument that has been expressed by many authors that the radiosensitivity (and possibly the volume dependence) of the esophagus varies along its length. In this case, different radiosensitivity parameters should be used in different radiotherapy sites to associate the delivered dose with the expected clinical outcome. It should be mentioned that most of the patients that were treated with the unilateral technique had no complications. This is strong evidence that proximal esophagus is a relatively parallel organ being characterized by a high volume dependence. The observed relative seriality can be influenced by the reference length or volume of the organ examined. If a very small part of the esophagus receiving uniform dose had been selected instead, then the relative seriality would appear to be very low. This is because it would not be possible to differentiate the response of the individual FSUs and identify their association since all of them would receive very similar doses, having consequently very similar response probabilities. The best way to observe the true relative seriality is by selecting a large part of the organ and utilizing irradiation techniques that produce significant dose degradation inside the organ. In this way, the volume effect can be estimated since the sparing of part of the organ will clearly show if this is adequate to retain the function of the organ.

The irradiation of the complete circumference of the esophagus is more likely to produce a stricture in comparison to a partial irradiation. This is denoted in our results by the low relative seriality (s) that was observed. An advantageous feature of our material is that the plane dose distributions shown in the upper graphs of Figs. 1–3 characterize almost all the planes up to 5 cm, which is the reference length used. Otherwise the planes that are completely irradiated should be examined in a different way than those that partially receive the dose. In this case, the association of the exact location of the stricture, with the dose delivered precisely, would be needed rather than selection of a longer length of esophagus. In this analysis, there was no evidence on the effects of radiotherapy on the Meissners and Mynteric plexes or the impact of the anatomical variations in the esophageal structure along its length. Nevertheless, such kind of information has been included in the new routine registration protocol that will soon be used in the clinic. During our analysis we investigated several factors that may affect the formation of esophageal strictures in the region of the proximal esophagus. As shown in Table 1, the location of the tumors varied and the large majority of them were not in the vicinity of the proximal esophagus to introduce significant structural changes to the tissue. It seems that as we move from stage T1 to T4 it is the treatment methodology that changes significantly, requiring a very different dose distribution delivered to the region. This dramatic change in dose is indicated to be the major factor for the esophageal strictures, making T3–T4 tumors more associated with the manifestation of this complication.

The quantification of the dose-response curve uncertainty as a function of the parameter uncertainties was calculated from the whole hyper-volume of the log-likelihood space, and not only from a selected plane in the parameter space. It constitutes primarily a qualitative evaluation of the effect of including parameter uncertainties

in the radiobiological model analysis, showing that the final result of the fitting process is a dose-response curve within an interval. The inclusion of the uncertainties characterizing the dose-response curves is important for the introduction of radiobiological modeling into the clinical routine.

The statistics in Table 3 have the purpose of showing that the estimated parameters lead to a very good agreement between the predicted complications and the true complication incidences for the study population, which to some extent is expected since the estimation of the model parameters was done using the study population. However, they also serve the purpose of recommendation as a means to check if another patient population (meaning another irradiation technique or treatment methodology) is compatible with the derived model parameters. This is something that should always be done before using parameters that have not been derived from the clinic, with the aim also that local data will be derived to support the validity of the initially used parameters.

The present analysis and the dosimetric characteristics of the study group could not associate any particular length or segmental relationship of the irradiated proximal esophagus with the expression of esophageal strictures. In other words, the observations of this study do not help us identify whether or not different parts of esophagus have different radiosensitivities. The clinical implications could be important regarding the factors affecting the expression of esophageal strictures. Furthermore, the size of our sample is too limited to draw such kinds of conclusions. Other important risk factors, such as bacterial infections and simultaneous surgical treatments, should be included in a more stringent analysis.

The NTCP calculations used in the present work are all based on esophageal DVHs, which were considered to bear the complete dosimetric information since they were calculated from the 3D dose distributions. The use of a scalar quantity, such as the mean dose, to quantify normal tissue response could of course simplify the calculations but at the cost of losing much of the structure in the data. The very inhomogeneous esophagus dose distributions characterizing our data set, due to the different treatment techniques applied, do not allow this approximation since it is only valid under the hypothesis of having small dose variations. The estimated radiobiological parameters differ in value from corresponding published data, which were calculated from other parts of the esophagus than the one examined in this study. This agrees with the viewpoint of many radiation oncologists that the radiosensitivity of the esophagus is not uniform over its length. As a consequence of these results, not only dose and irradiated subvolumes but also the location of the subvolumes could become key factors in NTCP modeling. In the light of these observations the data set studied in the present analysis could be further analyzed by using a model, which would take into account the location of the different partial volumes irradiated.

The radiobiological evaluation of the treatment plans may allow prediction of complications taking into consideration the variation of inter-patient radiosensitivity. This variation is partly expressed through the confidence intervals calculated in this study. The parallel presentation of the radiobiological data together with the dose distributions shows their close association. The use of radiobiological parameters is necessary if a clinically relevant quantification of a dose plan is needed. These parameters incorporate into the treatment plan evaluation the biological (clinical) information, which is needed to relate the dose delivered to a patient with the clinical findings that will follow. However, introduction of such parameters into the clinical routine requires care.

CONCLUSIONS

The present data set, consisting of individual complication data and DVHs for 82 patients irradiated for head and neck cancer, could be fitted by the relative seriality model. The application of the appropriate fractionation correction on the dose volume histograms before using them can be important. The dose-response curve of radiation-induced esophageal stricture is described by the parameters $D_{50} = 68.4$ Gy, $\gamma = 6.55$ and $s = 0.22$ for the reference length of the upper 5 cm of the proximal esophagus. The estimated radiobiological parameters show that the probability of stricture induction after radiation therapy is volume-dependent. The impact of the parameter uncertainties, which were quantified from the hyper-volume of the parameter space, on the dose-response curve was considerable. The goodness of fit was evaluated by three different statistical methods. The clinical utility of the determined model parameters was illustrated by applying them on a subset of the study material (probability of finding a worse fit = 61.0%, area under the ROC curve = 0.84 and χ^2 test = 0.95).

The present study provides much better information regarding the association of esophageal stricture formation with the dose volume information applied to the patients treated for head and neck cancer. Previous such information is very limited for the proximal esophagus and the present study is of the first to examine these cases using radiobiological modeling. Further independent studies using the same endpoint need to be compared. For the radiation-induced esophageal strictures, investigation of the surgical and bacterial infection history of the patients and the effect of simultaneous operative treatments could be very useful.

APPENDIX

Direct maximum-likelihood calculations are suitable for two particular types of problems: (i) low-statistics experiments with insufficient data to satisfy the requirement of

Gaussian statistics for individual histogram bins; and (ii) experiments in which the fitting function corresponds to a different probability density function for each measured event (meaning that each patient has a different weight in the calculations which is determined by the different dose distribution received) so that grouping the data leads to a reduction in information and loss of sensitivity in determining the parameters. Since it is not possible to extract more than minimal information from a very small data set, it should be expected that the direct maximum likelihood method is most useful for intermediate problems with modest data samples.

In this work, the fitting of the clinical data to the response parameters is done using the maximum likelihood method. To apply this procedure, it is assumed that there is a function that predicts the radiation-induced complications (the relative seriality model in this case). Such functions generally consist of two types of parameters. One type is related to the radiobiological characteristics of the organ under study describing its radiosensitivity, $\vec{X}$ and is usually model dependent (for the biological model that is used in this work these parameters are the D_{50}, γ and s). The other type of parameter is related to the characteristics of the treatment applied on each patient $\vec{\theta}$ (in this case the set of treatment variables are the doses and the corresponding subvolumes of the organ receiving these doses $(\vec{D},\vec{V})$). The treatment outcome of the patients, meaning the pattern of clinical findings of the patients participating in the fitting, is represented by two separate terms. The first term corresponds to the patients that had complications because of the treatment and the second term corresponds to the patients that were free of symptoms. The maximum likelihood function L, is then defined as:

$$\begin{aligned} L(\vec{X}|\vec{\theta}) &= L((D_{50},\ \gamma,\ s),(\vec{D},\ \vec{V})) \\ &= \prod_{i=1}^{m} P_{\mathrm{I}}((D_{50},\ \gamma,\ s),(\vec{D}_i,\ \vec{V}_i)) \\ &\quad \times \prod_{j=1}^{n}(1 - P_{\mathrm{I}}((D_{50},\ \gamma,\ s),(\vec{D}_j,\ \vec{V}_j))), \end{aligned} \quad [A1]$$

where the indices m and n run over the patients with and without complications respectively. The first product is the multiplication of the complication probabilities predicted by the radiobiological model for each patient with stricture. In the perfect situation, all these probabilities would be equal to 1 since all these patients develop the complication. In practice, the individual probabilities and their product are lower than 1. The second product concerns patients without complications. In the perfect situation their predicted complication probabilities would be equal to 0 so that the product of the second term would be equal to 1. In practice, these probabilities are larger than 0 and the second product is also lower than 1. In other words, the overall product gives the deviation of the individual complication predictions from the real clinical results. For computational simplicity and efficiency, the calculations were then performed using the logarithm of the likelihood function, $\ln L$:

$$\ln L((D_{50},\ \gamma,\ s),(\vec{D},\ \vec{V})) = \quad [A2]$$

The logarithm of a product is equal to the sum of the logarithms of the individual terms (e.g. $\ln(P_1 \cdot P_2) = \ln P_1 + \ln P_2$).

$$\sum_{i=1}^{m} \ln(P_{\mathrm{I}}((D_{50},\ \gamma,\ s),(\vec{D}_i,\ \vec{V}_i))) + \sum_{j=1}^{n} \ln(1 - P_{\mathrm{I}}((D_{50},\ \gamma,\ s),(\vec{D}_j,\ \vec{V}_j))) = \quad [A3]$$

Substituting the response function $P_{\mathrm{I}}((D_{50},\ \gamma,\ s),(\vec{D},\ \vec{V}))$ with the mathematical expression of the relative seriality model that is given by the combination of Eqs. [1] and [2], the following formula is obtained:

$$\sum_{i=1}^{m} \ln\left(\left[1 - \prod_{k=1}^{M_i}(1 - (\exp(-e^{e\gamma - (D_{ik}/D_{50})(e\gamma - \ln(\ln(2)))})^s)^{V_{ik}}\right]^{1/s}\right) + \sum_{j=1}^{n} \ln\left(1 - \left[1 - \prod_{k=1}^{M_j}(1 - (\exp(-e^{e\gamma - (D_{jk}/D_{50})(e\gamma - \ln(\ln(2)))})^s)^{V_{jk}}\right]^{1/s}\right) \quad [A4]$$

Eqs. [A1] and [A4] define in practice the probability that the biological model describes the observed pattern of complications, i.e. the probability that a group of $m+n$ patients, who receive the given dose distributions, has the observed outcome. The best estimate of the model parameters of the likelihood function, (namely the components of the vector $\vec{X}$), are the ones that maximize the likelihood function (or its logarithm) minimizing this way the deviation between the predicted and the clinical results.

Under the assumption that the errors of the estimated parameters are normally distributed, the variance-covariance matrix can be used to determine the confidence interval of any single parameter. Individual 68% confidence regions are given by the square root of the diagonal elements of this matrix. However, to find multi-dimensional confidence regions for non-normal or pathological likelihood functions one should search numerically for the values of the parameters for which the hypersurface (or hyper-volume) $\ln L(\vec{X}|\vec{\theta})$ intersects the surface (or volume) $\ln L = \ln L_{\max} - \frac{1}{2}\chi_a^2(k)$, where k is the number of dimensions of the likelihood function. The hyper-volume $\ln L(\vec{X}|\vec{\theta})$ is produced by calculating the values of the log-likelihood function for different values of $\vec{X}$, meaning D_{50}, γ and s. One then treats the region $[\vec{X}_{\min}, \vec{X}_{\max}]$ as a confidence region of probability content a. Most of the time, the probability a corresponds to the confidence regions of 68% or 95%.

Generally, there is not any convenient test of the quality of fit for the direct maximum likelihood method. The value of the maximum of the likelihood function itself is not useful, because it represents only the maximized probability

for obtaining the particular experimental result and there is no way of predicting the probability that should be expected. An estimate of the goodness of fit can be obtained by making a histogram of the observation data (clinical results) and comparing it to their theoretical predictions based on the best estimate of the parameters. Calculating the complication probabilities for each patient $\hat{P}_1^i = P_1^i(\hat{\vec{X}}|\vec{\theta}_i)$ using the best estimates of the model parameters $\hat{\vec{X}}$ and their dose distributions $\vec{\theta}_i$ and averaging over all patients, the mean value of $\ln L$ can be calculated as follows:

$$\overline{\ln L} = \sum_{i=1}^{m+n}[(1-\hat{P}_1^i)\ln(1-\hat{P}_1^i) + \hat{P}_1^i \ln(\hat{P}_1^i)], \qquad [A5]$$

and $V_{\ln L}$, which is the variance of $\overline{\ln L}$ is then given by:

$$V_{\ln L} = \sum_{i=1}^{m+n}\left[\hat{P}_1^i(1-\hat{P}_1^i)\left(\ln\left(\frac{\hat{P}_1^i}{1-\hat{P}_1^i}\right)\right)^2\right]. \qquad [A6]$$

Under the assumption that the value of $\ln L$ is normally distributed, Eqs. [A5] and [A6] can be used to determine the goodness of fit by comparing the mean and maximum values of the $\ln L$ considering the magnitude of the variance. From this comparison, it can be determined whether the probability of reaching a better fit between the predicted and the clinical results is high or not (margin of improvement).

ACKNOWLEDGEMENTS

The authors are grateful to Mats Beckman, Department of Radiology, Karolinska Hospital for his help in finding and selecting the post-irradiated radiographs shown in the work. This research was supported by grants from the Cancer Society in Stockholm, The King Gustaf V Jubilee Fund, Stockholm, and within the Center of Excellence by The Swedish National Board for Industrial and Technical Development.

REFERENCES

1. Emami B, Lyman J, Brown A, et al. Tolerance of normal tissue to therapeutic irradiation. Int J Radiat Oncol Biol Phys 1991; 21: 109–22.
2. Anseline PF, Lavery IC, Fazio VW, Jagelman DG, Weakley FL. Radiation injury of the rectum. Ann Surg 1981; 194: 716–24.
3. Roswit B. Complications of radiation therapy: The alimentary tract. Semin Roentgenol 1974; 9: 51–63.
4. Read PE, Ash DV, Thorogood J, Benson EA. Short-term morbidity and cosmesis following lumpectomy and radical radiotherapy for operable breast cancer. Clin Radiol 1987; 38: 371–3.
5. Sur RK, Kochhar R, Singh DP. Oral sucralfate in acute radiation oesophagitis. Acta Oncol 1994; 33: 61–3.
6. Teo P, Tai TH, Choy D, Tsui KH. A randomized study on palliative radiation therapy for inoperable non small cell carcinoma of the lung. Int J Radiat Oncol Biol Phys 1988; 14: 867–71.
7. Lind BK, Mavroidis P, Hyödynmaa S, Kappas C. Optimization of the dose level for a given treatment plan to maximize the complication free tumor cure. Acta Oncol 1999; 38: 787–98.
8. Mavroidis P, Kappas C, Lind BK. A computer program for evaluating the probability of complication-free tumor control incorporated in a commercial treatment planning system. J Balcan Union Oncol 1997; 3: 257–64.
9. Schultheiss TE, Oorton CG, Peck RA. Models in radiotherapy, Volume effects. Med Phys 1983; 10: 410–5.
10. Withers HR, Taylor JMG, Maciejewski B. Treatment volume and tissue tolerance. Int J Radiat Oncol Biol Phys 1988; 14: 751–9.
11. Wolbarst AB. Optimization of radiation therapy II: The critical-voxel model. Int J Radiat Oncol Biol Phys 1984; 10: 741–5.
12. Ågren A-K, Brahme A, Turesson I. Optimization of uncomplicated control for head and neck tumors. Int J Radiat Oncol Biol Phys 1990; 19: 1077–85.
13. Hendry JH, Thames HD. The tissue-rescuing unit. Br J Radiol 1986; 59: 628–30.
14. Dutreix J, Tubiana M, Dutreix A. An approach to the interpretation of clinical data on the tumor control probability-dose relationship. Radiother Oncol 1988; 11: 239–48.
15. Archambeau JO, Shymko RM. Tissue population configuration as a modifier of organ response. Int J Radiat Oncol Biol Phys 1988; 15: 727–34.
16. Söderström S. Radiobiologically based optimization of external beam radiotherapy techniques using a small number of fields. Thesis. Stockholm University, 1995.
17. Ågren A-K. Quantification of the response of heterogeneous tumors and organized normal tissues to fractionated radiotherapy. Thesis. Stockholm University, 1995.
18. Eriksson F, Gagliardi G, Liendberg A, et al. Long-term cardiac mortality following radiation therapy for Hodgkin's disease: analysis with the relative seriality model. Radiother Oncol 2000; 55: 153–62.
19. Gagliardi G, Björhle J, Lax I, et al. Radiation pneumonitis after breast cancer irradiation: Analysis of the complication probability using the relative seriality model. Int J Radiat Oncol Biol Phys 2000; 46: 373–81.
20. Källman P, Lind BK, Brahme A. An algorithm for maximizing the probability of complication free tumor control in radiation therapy. Phys Med Biol 1992; 37: 871–90.
21. Källman P, Ågren A-K, Brahme A. Tumor and normal tissue responses to fractionated non uniform dose delivery. Int J Radiat Biol 1992; 62: 249–62.
22. Brahme A. Treatment optimization using physical and biological objective functions. In: Smith A, ed. Radiation Therapy Physics. Berlin: Springer; 1995. p. 209–46.
23. Mavroidis P, Lind BK, Van Dijk J, et al. Comparison of conformal radiation therapy techniques within the dynamic radiotherapy project "DYNARAD". Phys Med Biol 2000; 45: 2459–81.
24. Mavroidis P, Theodorou K, Lefkopoulos D, et al. Prediction of AVM obliteration after stereotactic radiotherapy using radiobiological modelling. Phys Med Biol 2002; 47: 2471–94.
25. Aaltonen P, Brahme A, Lax I, et al. Specification of dose delivery in radiation therapy. Recommendations by the Nordic Association of Clinical Physics (*NACP*). Acta Oncol 1997; 10 (Suppl): 1–32.
26. Lind BK. Radiation therapy planning and optimization studied as inverse problems. Thesis. Stockholm University, 1991.
27. Lind BK, Källman P, Sundelin B, Brahme A. Optimal radiation beam profiles considering uncertainties in beam patient alignment. Acta Oncol 1993; 32: 331–42.
28. Löf J, Lind BK, Brahme A. Optimal radiation beam profiles considering the stochastic process of patient positioning in

fractionated radiation therapy. Inverse Problems 2003; 11: 1189–209.
29. Lepke RA, Libshitz HI. Radiation-induced injury of the esophagus. Radiology 1993; 148: 375–8.
30. Chabora BM, Hopfan S, Wittes R. Esophageal complications in the treatment of oat cell carcinoma with combined irradiation and chemotherapy. Radiology 1977; 123: 185–7.
31. Silvain C, Barrioz T, Besson I, et al. Treatment and long-term outcome of chronic radiation esophagitis after radiation therapy for head and neck tumors. Dig Dis Sci 1993; 38: 927–31.
32. Vanagunas A, Jacob P, Olinger E. Radiation-induced esophageal injury: A spectrum from esophagitis to cancer. American J Gastroenterol 1990; 85: 808–12.
33. Fowler JF. Brief summary of radiobiological principles in fractionated radiotherapy. Semin Radiat Oncol 1992; 2: 16–21.
34. Thames HD, Hendry JH. Radiation induced injury to tissues. In: Thames HD, Hendry JH, eds. Fractionation in radiotherapy. Philadelphia: Taylor & Francis; 1987. p. 1–21.
35. Herring DF. Methods for extracting dose-response curves from radiation therapy data, I: A unified approach. Int J Radiat Oncol Biol Phys 1980; 6: 225–32.
36. Jackson A, Ten Haken RK, Robertson JM, et al. Analysis of clinical complication data for radiation hepatitis using a parallel architecture model. Int J Radiat Oncol Biol Phys 1995; 31: 883–91.
37. Murtagh BA, Saunders MA. MINOS 5.4 User's guide, technical report SOL 83-20R, Department of operations research. Standford: Standford University, 1995.
38. Eadie WT, Drijard D, James FE, Roos M, Sadoulet B. Statistical methods in experimental physics. CERN: North-Holland Publishing Company, 1971.
39. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982; 143: 29–36.
40. Venkatraman ES, Begg C. A distribution-free procedure for comparing receiver operating characteristic curves from a paired experiment. Biometrika 1996; 83: 835–48.
41. Mavroidis P, Lind BK, Brahme A. Biologically effective uniform dose ($\bar{\bar{D}}$) for specification, report and comparison of dose response relations and treatment plans. Phys Med Biol 2001; 46: 2607–30.
42. Laurell G, Kraepelien T, Mavroidis P, et al. Stricture of the proximal esophagus in head and neck carcinoma patients after radiotherapy. Cancer 2003; 97: 1693–700.