I. Introduction

The aim of this chapter is to answer three questions, thereby defining the context of this thesis:

a. Why is accuracy in radiotherapy, and especially in conformal therapy treatment planning, crucial for advancements in treatment?

b. What accuracy is required in radiotherapy in general, and in treatment planning in particular?

c. How can compliance to established accuracy criteria be verified?

These questions correspond to section 1 through 3 of this chapter. Having set the context, the objective and outline of the thesis follow in section 4.

1. Conformal treatment, treatment planning and accuracy: conditions for advancements in radiotherapy

In this section the chain in dependencies is described between radiotherapy, conformal treatment, treatment planning, calculation accuracy and verification of this accuracy. It starts with a historical perspective.

1.1 Radiotherapy requires conformal treatment: a historical perspective

Already one year after the discovery of X-rays by Röntgen in 1895, a first attempt for radiotherapeutic treatment was made. The first documented cure by application of X-rays dates from 1899. Since then, increases in cure rate of patients with (mostly) malignant diseases treated with radiotherapy alone or in combination with surgery and/or chemotherapy have often concurred with improvements in radiotherapy treatment techniques. A milestone has been the introduction of cobalt-60 radiation and megavoltage radiation in the middle of the previous century. As shown in a recent overview, the new treatment techniques that became possible resulted in impressive increases in radiation treatment outcome in cases like prostate cancer, various head and neck cancers, uterine cervix cancers, as well as Hodgkin lymphoma. In the latter case, for example, the 10-years survival of M. Hodgkin patients inclined from 23% to 62%, and further increases have been reported since then. In the Netherlands approximately 50% of all cancer patients are presently cured, half of them by radiotherapy alone or by radiotherapy in combination with surgery. Further improvements are highly needed, as another estimated 8%-10% of all cancer patients is expected to benefit from improved local control. Common factor in all advances in treatment techniques is that they help to achieve the primary objective of radiotherapy: to deliver a sufficiently high and as homogeneous a dose as possible to the planning target volume (PTV), at the same time diminishing the chances for complications in surrounding tissue to the lowest possible level.

In the mid-eighties of the previous century, developments such as the introduction of the Beam’s-Eye-View projection tool facilitated the application of beams that much more conformed to the projected shape of the PTV as observed from the beam direction under consideration. Following the introduction of this tool the name ‘conformal radiotherapy’ became popular to stress the importance of a high-dose region just covering the PTV, avoiding critical structures as much as possible. Conformal radiotherapy has been defined as
1. Introduction

a procedure of high-precision irradiation of a target volume where the 95% isodose of the dose distribution, or more generally, the high-dose (treated) volume,\textsuperscript{10} conforms as closely as possible to the shape of the target volume in three dimensions. Whereas initially conformal radiotherapy was concerned with the optimal shape of radiation fields around the PTV, nowadays the focus is shifting to define also an optimal intensity distribution of energy fluence within the fields. This intensity-modulated radiotherapy (IMRT) opens the possibility to escalate dose at (parts) of the target with equal or lower complication chances, thus aiming at higher local control rate. A further step forward can be made by a combination of intensity modulation and limitation of the range of the radiation to the depth of the target, as is done in proton therapy.

Nearly a decade since the onset of this new generation of conformal techniques the first clinical results are showing up.\textsuperscript{11} Dose escalation in clinically localized prostate cancer has been reported to yield significant improvements. Initial clinical response in terms of prostate-specific antigen (PSA) levels was 90\% for patients that received 75.6 Gy or more minimum dose to the PTV, compared to 76\% and 56\% for patients that received 70.2 Gy and 64.8 Gy, respectively.\textsuperscript{12} Improvements were also reported in the rate of positive biopsies at 2.5 years and later after radiotherapy and in the 5-years PSA relapse-free survival.\textsuperscript{12,13} On the other hand, a dose escalation attempt in the nasopharynx did not show an improvement in treatment results.\textsuperscript{14} From a principal point of view, it is interesting to note that in these new techniques the objective of conformal treatment has been refined. The classical criterion of homogeneity of dose to the PTV is now explicitly balanced in the optimization process against criteria of maximum acceptable normal tissue dose. This extension of objective is sometimes indicated by the term conformal avoidance technique.\textsuperscript{15}

As shown here, the concurrency in improvements in radiotherapy treatment techniques and clinical outcomes can be expected to hold also for recently developed conformal treatments. A safe and effective application of conformal radiotherapy, and IMRT in particular, however, has profound consequences for treatment planning as well as for the required accuracy levels.

1.2 Conformal radiotherapy requires treatment planning

Treatment planning can be defined as the radiotherapy preparation process in which treatment strategies are defined in terms of planning target volumes, dose distributions tailored to these volumes, and sets of treatment instructions to deliver the dose distributions. Treatment planning plays a key-role in the advancement of radiotherapy. This key-role is underlined by the observation that four out of six historical suggestions for radiotherapy improvement as presented by Suit \textit{et al}.\textsuperscript{16} in 1988, are directly linked to treatment planning (a,b,e,f). Their shortlist of suggestions comprised: (a) better visualization of tumour and normal tissue, (b) presentation of uncertainty bands around isodoses, (c) monitoring of target tissue during treatment, (d) gating of treatment, (e) computer-controlled treatment and (f) reduced treatment volumes.

The present status of advances in treatment technique can be assessed in the light of this historical shortlist. The mentioned suggestions are now entering clinical practice. Modern treatment planning systems indeed have much improved visualization facilities (a), although especially accurate delineation of tumour and clinical target volumes remains to be one of the greatest challenges. In this perspective multi-modality imaging, such as CT, MR and/or PET
I. Introduction

images, and image registration\textsuperscript{17} can be of considerable value to better define and, if possible, reduce target volumes (f). Monitoring of targets (c) is readily approaching the realm of practical feasibility, based on markers of clinically acceptable size that can be automatically detected.\textsuperscript{18} Gating of a linear accelerator (linac) (d) triggered by the breathing cycle is an emerging application.\textsuperscript{19,20} The most evident example of computer-controlled treatment (e) is computer optimized IMRT, which can further improve conformation of the dose to the target volume with smaller fields and thus less normal tissue dose.\textsuperscript{21,22,23} When combined with protons as treatment modality, a further reduction in normal tissue dose is possible.\textsuperscript{24}

As indicated by these developments, treatment planning plays a key-role in the realization of enhanced conformal techniques.

1.3 Conformal treatment planning requires high and verified accuracy

The successful application of especially these enhanced conformal techniques depends critically on a very high level of accuracy in both treatment planning and treatment execution.\textsuperscript{25,26} The relevance of research into the accuracy of radiotherapy techniques and treatment planning follows from the observation, discussed in section 2 of this chapter, that optimal accuracy levels are currently mostly unachievable. Knowledge about the accuracy that is actually achieved is then of major importance. This applies similarly to the presentation of this accuracy.

Knowledge about the accuracy of treatment techniques is obtained by appropriate verification. This can be a verification of an entire technique, usually by application of a phantom in which calculated dose is compared to measurements, e.g. by film,\textsuperscript{27,28} a combination of scans in a water phantom\textsuperscript{28,29} and gel dosimetry.\textsuperscript{30} In case of proton treatment the use of a PET scanner for dose verification has also been suggested.\textsuperscript{31} Furthermore, a portal dose image can be calculated by the treatment planning system and compared to a measured portal dose image.\textsuperscript{32,33,34} Thus the overall accuracy of the relative or absolute dose at a particular point for a particular treatment technique is explored. The results are applicable to the investigated technique only. The disclosure of the causes of any discrepancies thus found then requires a subsequent ‘backward’ analysis to discriminate the constituting steps.

Verification can also be confined to an individual treatment step, to assess whether appropriate accuracy requirements for that step are complied to. It is evident that an appropriate combination of both scopes of verification, overall and step-oriented, is most powerful. Overall verification might point at weaknesses in the step-oriented verification. Also, when results of subsequent steps can be combined, discrepancies might be anticipated before their effects appear in overall treatment verification.

It is important to note that discrepancies in treatment planning calculations, even when relatively small in magnitude, constitute systematic errors in patient treatments and therefore require proper attention. An extensive verification program is needed to explore the presence and magnitude of such discrepancies. This is described in section 3 of this chapter.

The required knowledge of treatment (planning) accuracy implies the necessity to present the treatment (planning) accuracy in a proper way. This requirement is resembled by the second item of the above-mentioned shortlist, the presentation of uncertainty bands around isodoses (b) or, more generally formulated, the presentation of uncertainty in treatment planning.\textsuperscript{35}
Ideally, a comprehensive presentation of all treatment uncertainties should be part of the treatment planning process, preferably also expressed in radiobiological outcomes. It has been shown that inclusion of treatment uncertainties in treatment planning effects the outcome of treatment plan optimization and treatment plan robustness. Many studies have been dealing with organ motion and/or setup-error and with methods to account for these uncertainties in radiotherapy treatment planning. Also adaptation of a treatment plan based on portal image data is being considered. However, the implementation in routine treatment planning has up till now not been specified. Still further away is incorporation of uncertainty in radiobiological modelling, which is yet a rather novel area of research.

Surprisingly, the presentation of errors in treatment planning calculations in the treatment planning process itself is uncommon. Doing first things first, there is conceptual logic in an exploration of the uncertainties that originate in treatment planning itself, which is the objective of this thesis, and in a presentation of these uncertainties in treatment planning, which is discussed in chapter IX of this thesis.

Summarizing this introductory section, it has been shown that knowledge about the accuracy of treatment planning calculations is a crucial requirement for the enhancement of treatment planning capabilities, and thus an essential condition to achieve the expected benefits of conformal radiotherapy treatment. This thesis aims at the acquisition of such knowledge.

2. Accuracy requirements for treatment planning calculations

In this section a fundamental and a practical approach are followed to define requirements for calculation accuracy, and an attempt is presented to bridge the gap between these approaches.

2.1 Fundamental approach

Accuracy requirements for radiotherapy treatment should basically be derived from radiobiological behaviour of tumour cells and normal tissues, as well as from clinical evidence that represents this behaviour. In this context it should be kept in mind that clinical evidence obtained so far has been hampered by uncertainties in delivered and reported dose and variations in tumour cell characteristics, so that requirements might need to be stricter to resolve these effects in future. The latter effect has been studied by Webb and Nahum, who indicate a much steeper dose-effect curve and thus much higher dose accuracy demands if these variations would be taken into account.

The wide diversity in tumour sensitivities and normal tissue tolerances found in practice could lead to a range of accuracy requirements. As this is considered to be undesirable in practice, requirements should be based upon the most critical situations encountered in regular radiotherapy practice. Specific applications concerning, for example, single fraction radiosurgery of benign diseases can be excluded as these might demand facilities such as invasive stereotactic frames that are not compatible with fractionated radiotherapy. It is interesting to note in this respect that fractionated treatment with a relocatable frame is reportedly associated with a more accurate dose prediction than single fraction treatment with an invasive frame.
In a recent report, an overview is given of clinical evidence for the importance of radiotherapy accuracy, from which it is concluded that a difference in absorbed dose of 10% is detectable for tumours and that a difference of 7% in absorbed dose can sometimes be observed for normal tissue reactions. For the latter figure even 5% has been reported. From an extensive review of dose-response data, noticing that this concerned cell and population averages, Brahme et al. concluded that the standard deviation in mean dose should be at most 3% (1 S.D.) to have a reasonable control of the treatment outcome (with a 5% action level). This is in agreement with a review undertaken by Mijnheer et al. of dose-response steepness data observed in normal tissue complications. They concluded that transfer of these data between institutes requires the dose to be known at the specification point within 7% accuracy, which they equated to 2 S.D., noting that equating it to 3 S.D. would be not achievable, at least not at that time (1987). A critical interpretation of their most sensitive data, neglecting practical feasibility would thus result in an even stricter requirement, for example, 2% taken as 1 S.D., which has been suggested as the most critical requirement (no standard deviation defined) in ICRU Report 24. This is in reasonable agreement with a Nordic report that suggests a dose accuracy of just 1 Gy (1 S.D.) for steep responding tumours, corresponding to approximately 1.5% for common dose prescriptions.

Considerations of heterogeneity in dose and tumour cell characteristics over the target volume have been used as an argument for less strict dose requirements in other points of the target, leading to 5% (1 S.D.) over the entire target volume. However, an assumed spread in tumour cell densities will not allow for much spread in dose, because studies show that even large variations in tumour cell density only permit very modest variations in dose. Moreover, it is also noted in the same report that it is likely that a better accuracy will be needed when better tumour control data is obtained. Thus the proposed relaxation of requirements to 5% in other points than the specification point is questionable.

It must in addition be stressed that these figures apply to the uncertainty for the total radiotherapy treatment, so that requirements for treatment planning are necessarily stricter.

Another approach has been followed by Ahnesjö. He defined a practical limit where further increase in dose calculation accuracy does not yield an increase in total treatment accuracy, taken into account reported uncertainties in calibration and delivery technique. His conclusion is that at present there is no need for a dose calculation accuracy better than 2%, whereas this ultimately might be confined to 1%. Similarly, Mackie et al. reasoned that dose calculation accuracy need not be as accurate as absolute dose calibration while, on the other hand, calculated dose should not deviate too much from prescribed dose; they concluded that dose calculation accuracy should be in the range of 2% to 5%.

In the absence of unambiguous conclusions from tumour control studies, a criterion of 4 mm (1 S.D.) has been formulated on the position of beam edges. This criterion is stated to include all geometric and movement factors. In recent reviews standard deviations down to approximately 1 mm have been given for both organ motion and setup-error for some treatment sites and motion directions. Reasoning that treatment planning should not be the weakest link in the chain, a criterion of 2 mm (1 S.D.) or less seems appropriate for treatment planning alone. The stated total criterion of 4 mm may than be considered as too relaxed. Specific requirements such as leaf position in some dynamic IMRT fields need to be tighter, e.g. 1 mm.

Concluding the fundamental approach, accuracy criteria of 2% (1 S.D.) in calculated dose and 2 mm (1 S.D.) in field edge position can be seen as generally appropriate, whereas 1% or 1 mm might be required in specific circumstances.
2.2 Practical approach

As the accuracy levels derived from this fundamental approach can not be achieved at present, practical approaches are frequently considered. These are based on expert opinions and clinical experience combined with assessments of practicability, resulting in a wide variety of criteria for treatment planning accuracy. A comparative overview has been given by Venselaar et al., and some references have also been included in chapter III. Published criteria cannot in a straight-forward way be compared to each other due to differences in reference dose, local or absolute, isocentric or $d_{\text{max}}$, and to differences in situations in which the criteria should be applied. Venselaar et al. undertook a conversion of the most recent, elaborated criteria to four beam geometries in a phantom, and reported many differences between these criteria of up to a factor of 3 and incidently higher. Thus general consensus based on expert opinion and experience is still far away. Consensus is, however, found to a large extent in a high dose, low gradient region in a simple geometry in a homogeneous phantom, where many authors/reports define 2% in local dose as criterion, although others define criteria for this geometry in the range of 1% to 4%. Strictly, the latter is announced as a suggestion, not a requirement. Some of these references complement this requirement with a 2 mm limit in cases of high dose gradients, while others apply less strict criteria up to 4 mm.

In these practice-based approaches criteria are relaxed in cases where a planning system is probably less accurate or where inaccuracies accumulate. For instance, simple, complex and more complex or anthropomorphic phantom geometries, and central axis and off-axis criteria are distinguished, and correspondingly relaxed criteria are applied. Outside the beam, accuracy criteria of up to 5% relative to the central axis dose have been specified, whereas application of criteria from other reports to the same geometries may yield even higher values. This widening of tolerances thus resembles expected shortcomings in many present treatment planning systems. But it might be less appropriate for some other treatment planning systems, depending on the sophistication of the algorithms used. For the time being it can be seen as the best practical achievable solution, for example for treatment planning intercomparisons or to identify unexpected large errors for any particular treatment planning system. However, it is recognized that these (actual) tolerances should not reflect levels of acceptability beyond which no further improvements are necessary, but indeed should converge to the criteria derived from radiobiological and clinical data. Furthermore, the presented practice-based accuracy criteria might already be too tolerant in the assessment of advanced treatment techniques that aim at small, albeit possibly significant dose differences, for example, a reduction of 1-3 Gy at the 70 Gy dose level in the rectum during a prostate treatment.

If general limits are desired, values of 3% or 3 mm may be considered as representative. These tolerance levels were recently specified for a ‘complex geometry’ with wedge, inhomogeneity or asymmetry.

In the application of the criteria cited above, confusion can arise as to where the dose and where the distance criterion should be applied. A boundary at a dose gradient of 30%/cm has been proposed. On the other hand, the IPEM Report 68 states that this confusion should be clarified in a department policy statement. Although this yields clarity in an institute, it does not promote uniformity between institutes. Moreover, any choice for a boundary inevitably has inherent arbitrariness.
Concluding this practical approach, accuracy criteria were found to vary widely between reports/authors and between situations with more or less complexity. Especially in more complex geometries a clear gap exists between the criteria set by the practical approach and those defined by the fundamental approach. Furthermore no unambiguous transition exists between dose and distance criteria.

2.3 Bridging the gaps

The two gaps that have been sketched in the previous section can (at present) not be eliminated. The first, the different criteria resulting from the fundamental and the practical approach, will exist until treatment planning has further evolved to let the criteria of both approaches converge. The second, the transition between application of dose and distance criteria, inherently depends on a subjective assessment as to what criterion has to be applied at any specific location. In practice, it is experienced that the existence of these gaps diminishes the efficacy of treatment planning verification.

Faced with their existence, a solution is sought in this thesis to bridge the gaps, namely by application of a function that combines distance and dose discrepancies in an intuitive way, prior to applying any criterion. Such a function is introduced in chapter III. This so-called field accuracy, Fig. 1, is defined for comparison of calculated and measured dose distributions as the difference in percent or the distance-to-agreement in millimetres, whichever is less. The

Fig. 1 Illustration of the field accuracy concept. A calculated (thin line) and a measured (thick line) schematic dose profile are shown in (a). In (b) the difference in dose between calculation and measurement is shown as a solid line. The distance-to-agreement of calculation and measurement, carrying the ± sign of the dose difference, is shown as a dashed line. The corresponding field accuracy is shown in (c). Regions where the field accuracy curves are determined by mm-curve distance are thickened. Crossings of dashed horizontal lines indicate that an accuracy criterion, here 3 [%;mm], is exceeded.
combined pseudo-unity is indicated as [%;mm]. The intuitive element in this definition is that use is made of the observation that distance and dose criteria in current units are numerically close to each other, in many practical (near 3% and 3 mm) and ideal (near 2% and 2 mm) criteria. A similar approach has been proposed by Harms et al. The field accuracy data is not \textit{a priori} reduced to a single ‘goodness’ quantity, but the accuracy level is verified \textit{a posteriori}, thus maintaining the possibility to compare to any criterion, e.g. 2 or 3 [%;mm], as long as the assumed numerical equivalence of dose and distance holds. In case of large amounts of verification data, which is common practice in a full dosimetric verification of a modern treatment planning system, the use of this quantity should preferably be embedded in a practical, dedicated procedure, which is presented in section 3. The application of this field accuracy concept has markedly improved the efficacy of the verification procedure.

A further advantage of the application of this field accuracy concept is that no concession to accuracy requirements is accepted based on assumed computational weaknesses of a treatment planning system’s algorithm. On the contrary, it is considered that in those cases (treatment techniques) where less calculation accuracy is found, this should just be presented as such, and then balanced against the clinical advantage, if any, of the application of such a treatment technique. Such an error documentation without \textit{a priori} concessions should be an integral part of the verification data, section 3. Preferably the field accuracy presentation should also be available in combination with any (kind of) treatment plan, to give an impression of the overall (un)acceptability of treatment planning accuracies in targets and tissues in such a treatment plan. The feasibility of this idea is discussed in Chapter IX.

A critical note to all quantities representing geometric accuracy, including field accuracy, is that ultimately not distance-to-agreement but tumour control probability is the most relevant quantity. But even when this would become an integrated element of treatment planning, it is expected that verification practice will always also need handsome geometric criteria.

In conclusion, accuracy criteria for treatment planning calculations would ideally be 2% or 2 mm, and 1% or 1 mm in specific circumstances, whereas 3% or 3 mm is often seen as presently acceptable. The field accuracy concept allows an adequate comparison to the combined criteria.

### 3. Verification of treatment planning dose calculations

A verification programme (procedure) is required to check compliance to established accuracy criteria. The extent of this programme and its implementation into a program (software) are described in this section. The power of the program is illustrated by an example.

#### 3.1 Combinatorial verification

The high level of treatment planning accuracy described above necessitates a well-structured verification programme. The structure of such a verification of a treatment planning system can be understood by the analogon of the defences of a medieval city based on several concentric walls Chapter II. In case the outer wall fails, the inner walls will hopefully stand up. The outer wall in treatment planning verification is the verification work (release test) provided by the manufacturer. The scope of the present section is the next inner wall, the in-
hospital verification (acceptance procedure) of a new treatment planning system, a new algorithm or just a new release of an existing system or algorithm. A further inward wall is represented by system quality control such as checks that no changes have occurred in a system which has been released for clinical use, and checks of, for instance, the correctness of Hounsfield units of a CT scanner. The most inner wall is patient quality control such as a check on monitor units in every treatment field of any patient.

In recent years many reports have been published about treatment planning verification. These reports provide tests and procedures for functional, geometric and dosimetric verification. This section is confined to verification in the sense of dosimetric acceptance testing aimed at defining the level of accuracy of dose calculations in a treatment planning system, and at identification of treatment techniques where sufficient accuracy would not be warranted. (It must be noted that dosimetric and geometric verification cannot be separated completely, because dose in the penumbra is directly linked to the definition of beam edges, as will be shown in section 3.3.) The suggestions in the mentioned reports provide a basis for such an acceptance test, a basis that has to be adapted to tailor it to the characteristics of the local treatment planning system and possibly also to meet the requirements of specific clinical applications. The model in our treatment planning system (Helax-TMS, MDS Nordion, Canada) is an energy-fluence based model. The tailoring of the verification program to the characteristics of this model is described in Chapter II.

The number of checks that has to be performed increases rapidly with the degrees of freedom of a treatment planning system and the variety of desired clinical applications. A typical set of parameters for water phantom scans used in our hospital for the photon beams of a triple energy linac equipped with MLC is given in Table I. This set is to be extended with situations of oblique incidence, missing tissue and tissue heterogeneities.

Table I. A typical set of parameters for water phantom scans used in the acceptance test of the photon beams of a new linac for dose calculation in a treatment planning system.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Values</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy</td>
<td>6, 10, 15 MV-X</td>
<td>3</td>
</tr>
<tr>
<td>Source-surface distance (SSD)</td>
<td>75, 90, 120 cm</td>
<td>3</td>
</tr>
<tr>
<td>Field sizes</td>
<td>3, 5, 10, 15, 20, 30, 40 cm²</td>
<td>7</td>
</tr>
<tr>
<td>Field shape</td>
<td>square, elongated, asymmetric (half/quarter)</td>
<td>3</td>
</tr>
<tr>
<td>Field edge</td>
<td>collimator, block, MLC</td>
<td>3</td>
</tr>
<tr>
<td>Modulation</td>
<td>open, wedge</td>
<td>2</td>
</tr>
<tr>
<td>Scan direction</td>
<td>X, Y, Z</td>
<td>3</td>
</tr>
<tr>
<td>Depths</td>
<td>(d_{\text{max}}), 5, 10, 20 cm</td>
<td>4</td>
</tr>
</tbody>
</table>

Although the number of parameters in this example is modest, merely representing regular use of a modern linac, the number of combinations of these parameters is immense, corresponding to over 10,000 scans. This resembles the observation that each separate parameter might be interpreted as an extra dimension spanning a huge ‘verification space’. Of course, an intelligent reduction of this number is well possible, because not all ‘grid points’ in verification space need to be explored, but definitely much more has to be done than a straight-forward exploration of the main axes of verification space. For instance, it might well
be that calculations in open fields, and fields supplied with a shielding block or a wedge are all in good agreement to measurements, however, this gives no guarantee that in a combination like a wedged block field the dose is calculated with acceptable accuracy. Moreover, the calculation of such a combination might happen to constitute the decisive element in a clinical evaluation, for example, if rectum dose becomes the decisive criterion in a prostate treatment.

In Fig. 2 the combination of wedge and blocks is illustrated in some verification results of our treatment planning system, Helax-TMS 5.1. Measurements have been performed with a ionization chamber (IC15, Wellhoefer, Germany) in a 15 MV-X beam in a water phantom (Blue phantom, Wellhoefer, Germany). In combination with measurements in a reference field (10x10 cm², depth 10 cm at SSD=90 cm) absolute dose values are obtained, only excluding variations in linac output. Calculations also relate to absolute dose, so that absolute dose discrepancies, relative to 100% reference dose, are shown. In terms of a ‘classical’ treatment planning system, errors in output, depth dose curves, off-axis ratios and block transmissions are thus all included. The block used is an arrow-shaped shielding block, identical to that used in chapter III. For clarity of display only scans at 10 cm depth in a scan direction defined by the lower jaws are shown.

The lower curves in Fig. 2 represent the field accuracy. The field accuracy of the open field and the fields supplied with a wedge or a block, Fig. 2a, b and c, respectively, are mostly within ±3 [%;mm]. Under the shielding block the dose is overestimated by approximately 2.5%. In this case this slight overestimation is on the ‘safe side’, and probably of no clinical meaning. In contrast, the combination of block and wedge, Fig. 2d, shows an underestimation

---

**Fig. 2** Calculated (dashed lines) and measured (solid lines) line dose profiles (upper part of each panel) at 10 cm depth and SSD=90 cm in 15 MV-X fields: (a) open 20x20 cm², (b) wedge 20x20 cm², (c) blocked 20x22 cm² and (d) blocked wedge 20x22 cm². The corresponding field accuracy curves are plotted in the lower panels of (a), (b), (c) and (d), respectively. The inserts represent the applied field shape, wedge direction, if applicable, and the position of the line dose profile.
of the dose under the block. At the thin side of the wedge the underestimation amounts to 8% of the dose to the centre of the field. This underestimation was found to be confined to the highest photon energy (15 MV-X) and appeared to increase with decreasing SSD.

This example illustrates that ‘combinatorial verification’, defined as exploration of all relevant combinations of field parameters, is thus needed to discover situations of possibly unacceptable inaccuracy. Illustrated by a metaphor, combinatorial verification is required to find ‘dust traps’ that might otherwise remain hidden in the corners of parameter space.

In conclusion, combinatorial verification is required to check compliance to established accuracy criteria. As a consequence, a multitude of geometries has to be verified. The field accuracy concept is helpful to get a better overview of the results, but only with an adequate implementation of verification procedures this undertaking can be successful.

### 3.2 Implementation of combinatorial verification

The feasibility of combinatorial verification depends on practical aspects of implementation, concerning first of all the large amounts of data to be handled. For example, after thoughtful reduction of the set of parameters listed for one new linac in Table I, still approximately 1500 line dose profiles remained which had to be measured, in combination with approximately 200 beam output measurements to obtain absolute dose values. A thimble ionization chamber (IC15, Wellhoefer, Germany) has been selected for this purpose, because of its reputation of uncomplicated use, and the virtual absence of energy dependence and negligible direct photon response. An array of such ion chambers is a good equivalent to reduce the required linac time.76 With this type of dosimeter a correction must be performed for the detector’s non-ideal spatial behaviour by application of the line spread profile, chapter VI. In addition, a correction for small misalignments (assumed to be less or equal ± 1 mm) in the setup of the water phantom might be required, and a transformation to account for a difference in coordinate systems of water phantom and treatment planning system. The same amount of line dose profiles were calculated by the treatment planning system, which had to be connected correctly to the corresponding measured profiles and output factors. Comparison results should include a field accuracy profile, and for consistency checks also the field size, the penumbra steepness, and the manipulations performed on the data should be registered. The results of the comparisons should be displayed, documented and archived in an easily appreciable way.

To achieve this purpose a profile comparison program and a profile database were developed, as schematically represented in Fig. 3. A macro defines all actions to be performed thus preventing possible errors caused by direct user interaction. The results are documented in a library of comparison results with corresponding plots. Subsequently, the comparison results can easily be used for further analysis.

This verification facility is now in use at our hospital since 1995. At present our profile database contains approximately 3500 measured profiles which have been compared to corresponding calculations of the current version of the treatment planning system. The following results were obtained by this program so far:

- An appropriate verification of the calculations of an advanced treatment planning system has been performed.
- This verification is easily repeated when modifications in hardware (new linac) or software (new releases) require to do so.
• Errors in the calculation algorithm have been identified, leading to code analysis and bug fixes in new releases. Dose errors down to 1% could thus be traced. For instance, scatter from a part of a (small) wedge that was outside the projected maximum wedge field size but within the beam cone defined by the primary collimator was found to be omitted. This resulted in a local 1% dose error in the calculated output of wedge fields.

• Errors in linac field size were found in measurements that were made for the purpose of a beam-fit for the treatment planning system, Chapter II. For instance, a 1.5 mm positional error in the collimator position of a wedge field was identified in this way. This influenced the gradient of the calculated wedge profiles, causing a local 1.5% error in dose.

In conclusion, the implementation of combinatorial verification into a profile comparison facility and a profile database, as shown in Fig. 3, has proven to allow verification of treatment planning calculations over all required dimensions of ‘verification space’ and has contributed to improvements in both the calculation algorithm and the dosimetry for beam commissioning in the treatment planning system.

3.3 An example of verification analysis: field sizes

The data that result from the combinatorial verification procedure are fed into the database. This database is well suited to perform more advanced analysis, as is shown in this section. An extended application of this verification data is described in Chapter IX.
In Fig. 4 an example of analysis results is shown. In this figure differences in field sizes (widths) are shown that were found in a 15 MV-X beam (SLi15, Elekta, Crawley) at 90 cm SSD. The presented field sizes, in this figure indicated by W, are defined by the lower X-diaphragm, which is a pair of focussed collimator jaws. In Fig. 4a the differences in field size at 50% of the profile centre value of calculated and measured profiles are shown for symmetric field sizes of 5x5 cm$^2$ to 30x30 cm$^2$, at depths of 5 cm, 10 cm (isocentre) and 20 cm. The calculations were done with a pencil beam model in which no correction for off-axis softening of the energy spectrum is applied. The calculated field size at 50% height is seen to be systematically larger than the corresponding measured field size. Differences of up to 1.3 mm were found, with the greater differences found in larger field sizes and at greater depths. The deviations are less than 1.0 mm for field sizes up to 20 cm, and the systematic character might have passed unnoticed with normal comparison of profiles. Nevertheless, as this difference occurs in between all treatment plans and treatment executions a closer inspection is appropriate.

Fig. 4 Field size differences in dose profiles of a 15 MV-X beam at SSD=90 cm at depths of 5 cm (triangle), 10 cm (circles) and 20 cm (squares). (a) Differences between calculated and measured field sizes, defined at 50% of the profile centre value. (b) Differences of expected (=geometrical) field size and measured field size (at 50% height). The dashed lines indicate estimated differences at 5 cm and 20 cm depth based on measurements at 10 cm depth. (c) Schematic dose profile (error function) of a single source (thin solid line) and combination of two sources where the second source yields an extended 10% dose contribution (thick solid line). The geometric field edge position is indicated by a vertical line. In the single source geometry the positions of 50%-points and inflection points coincide with the field edge. In the dual source geometry the 50% points are moved outwards. (d) Difference of geometrical field size and measured field size, as determined by the position of the inflection points on the profile.
Further analysis yielded two causes for these deviations and one other was excluded. As the mean difference over the range of field sizes between measured (50%) and expected (=geometrical) field size at isocentric depth was 0.0 mm, Fig. 4b, an error in the regular calibration procedure of the jaw positions could be excluded. Both at greater and smaller depths slightly greater differences are found especially in larger fields. Based on the differences at 10 cm depth and estimating the effect of off-axis softening on the position of the 50% points at other depths, the differences indicated by dashed lines in Fig. 4b were calculated, which mostly agree very well with the observed differences. This illustrates that calibration of collimator position is strictly speaking only valid at one depth. This first cause, field-size reduction at greater depths and off-axis position due to off-axis softening is not taken into account in the pencil beam calculations, causing the fan out of lines in Fig. 4a.

The second cause was analysed to be linked to the average difference of 0.4 mm that was found between the calculated width at 50% height and the geometrical field size. This phenomenon is schematically explained in Fig. 4c. A collimator field edge originating from a single source, such as the focus of a linac, is represented here by a sigmoid-like shape (thin line) where the position at 50% height coincides with the inflection point of the profile and with the geometrical field edge, i.e. the projection of the focus along the collimator edge. However, additional fluence that emerges from a second source, located lower in the linac, such as head scatter from the flattening filter, is associated with a broader field and thus effectively raises locally the dose profile that originates from the linac focus. Therefore, while the X-position of the inflection point remains almost fixed, the position that indicates height at 50% is moved slightly outwards. The only physical correct correlation is thus between geometric field size and inflection points. This correlation was checked for our treatment planning system and found to be correct within 0.010± 0.013 mm over the presented range of data. Ideally the calibration of the collimator position would use the inflection points. This is illustrated by Fig. 4d, where the differences between the geometrical and measured field sizes based on inflection points are shown. The close resemblance of these curves with Fig. 4a indicates that this difference is indeed the main cause of the differences found in Fig. 4a. However, a calibration at the 50% points has clear practical reasons. An additional correction of 0.4 mm is then a good remedy. For a 6 MV-X beam also a correction of 0.4 mm was found. It should be noted that these data are based on measurements in a water phantom, so that they may not directly be applicable to film measurements, in particular at depths beyond dose maximum where energy effects may play a role.

An approach to account for these effects could be to apply a water phantom for the collimator calibration, to define the depth at which the settings should be optimal (10 cm depth is considered to be a good choice), to calibrate collimator settings at positions of 50% height, and to add an extra 0.4 mm field width (if 10 cm depth was chosen) afterwards. The remaining maximum discrepancy in this set of data is than 0.3 mm at the depth of calibration and 0.9 mm elsewhere. Note that in this approach the calibration of the field sizes at the linac is adapted to a geometric beam definition, rather than the common 50%-based definition. However, a geometric beam definition is required to allow proper calculation of the contribution of multiple sources of linac energy fluence. The fan out of lines in Fig. 4a is expected to disappear when off-axis softening is taken into account in the calculations.

In conclusion, the described implementation of combinatorial verification is shown to provide a powerful tool for further analysis of beam data. In the example given a systematic difference in mean width of 0.4 mm over all field sizes was found between measured and
To define the accuracy of dose and volume calculations that are essential for conformal radiotherapy treatment planning.

4. Objective and outline of this thesis

In the previous sections it has been shown that advancements in radiotherapy, especially in conformal treatments, require that the accuracy of treatment planning calculations lays within margins that can be derived with reasonable approximations from radiobiological and clinical data. An extensive verification programme is necessary to check compliance of treatment planning calculations with these accuracy limits. The accuracy of treatment planning calculations can only be determined properly when the physics and mathematics involved in all calculation and verification steps is understood. This thesis is intended to gain such knowledge for calculations that are of fundamental importance for conformal radiotherapy treatment planning. Thus, the objective of this thesis is:

The papers included in chapters II - VII meet various aspects of this objective.

Chapter II describes the verification programme that has been designed to check the dose calculation accuracy in the treatment planning system that is applied at the Groningen University Hospital. This treatment planning system is based on an energy transport model. In this chapter the required tailoring of the verification programme to the physical characteristics of the calculation model is presented.

Chapter III is concerned with the first step, also in historic perspective, in conformal radiotherapy: the confinement of treatment fields to the projected shape of the PTV by individual shielding blocks. The accuracy of dose distributions calculated with pencil beam kernels in blocked photon fields is verified in comparison to measurements in a water phantom. In this chapter the concept of field accuracy is also introduced and shown to be a valuable verification tool.

Chapter IV focusses at one of the major challenges in treatment planning calculations: the calculation of dose in and near tissue heterogeneities. Due to extended transport ranges of recoiled electrons and scattered photons the penumbras of photon beams are considerably widened in low density materials such as lung, in comparison to more water-like tissues like muscle and adipose. This yields a lower dose near the edges of a beam and thus a possible too low dose in, for instance, the PTV of a lung tumour. In this chapter dose calculations have been verified by measurements in a phantom that simulates a lung tumour geometry. The calculations were performed by several algorithms that are currently in use in modern treatment planning systems, amongst others the pencil beam kernel algorithm with one-dimensional convolution correction algorithm described in Chapter II. Recently, the calculations were repeated in the same geometry with a point spread kernel algorithm, of which results are included in chapter II.

Chapters V and VI are dedicated to the accurate determination of beam edges. The conformance of a radiotherapy treatment depends on an accurate knowledge of the beam edges. Thus accurate measurements of beam penumbras are required to customize and verify
the calculation model of a treatment planning system. As outlined above, a large amount of measurements is required for combinatorial verification. Therefore, well-established routine dosimeters such as thimble ionization chambers and photon diodes are often preferred. The non-ideal behaviour must than be corrected for. This behaviour originates from the detector dimensions, and in case of ionization chambers, the alteration of electron transport and the replacement of medium by the materials of chamber wall and electrode. In chapter V analytical calculations are described that model the transport of Compton recoil electrons. The model is verified with measurements using a diamond detector in a telescopic slit beam geometry to be able to describe primary dose profiles of an elementary slit x-ray beam. In chapter VI this model is extended to calculate detector response profiles of a thimble ionization chamber. Calculated detector response profiles of ionization chambers and a photon diode are experimentally verified in the telescopic slit-beam geometry. The combination of detector response profiles and dose profiles yields the line spread function of this type of detector.

Chapter VII is devoted to accuracy in proton beams. When photons are replaced by protons as treatment modality, a further improvement in conformal radiotherapy can be achieved, primarily because of the well-defined depth of penetration of protons. In addition, very sharp lateral beam edges can be achieved with proton beams, especially close to the final beam collimation. Collimator scatter becomes important at short distances. Its contribution has been calculated with a Monte Carlo model. The calculations were verified by means of a two-dimensional detector consisting of a scintillating screen and a CCD camera. The line spread function of the detector was verified with a diamond detector.

Chapter VIII ends this series of papers by studying accuracy of volume calculations, the step logically following dose calculations in treatment planning. The three-dimensional dose distributions that result from treatment planning calculations are difficult to appreciate for a human observer, for example to assess whether a distribution complies with dose and volume criteria for targets and critical structures. Dose volume histograms are then a valuable tool. To obtain a dose volume histogram, a sampling of dose over the volume must be performed, usually based on a regular grid of dose points. The accuracy of this grid-based volume sampling has been the subject of some controversy. This controversy originated from a neglect of the influence of the shape of a structure on the accuracy of the volume sampling, as is proved in chapter VIII.

Chapter IX, a future perspective, investigates the feasibility of a presentation of calculation accuracy in a complete treatment plan. The importance of such a presentation follows from the observation that any discrepancy in the calculated dose distribution constitutes a systematic error in the treatment, and its magnitude should thus be known during treatment planning. This presentation provides the links between the research conducted in this thesis, the results of a verification programme on a treatment planning system’s calculation accuracy, and a presentation of calculation accuracy in individual treatment plans to help treatment decisions.
References

I. Introduction

I. Introduction


I. Introduction


