GUIDELINES FOR THE VERIFICATION OF IMRT
GUIDELINES FOR THE VERIFICATION OF IMRT

Markus Alber
Sara Broggi
Carlos De Wagter
Ines Eichwurzel
Per Engström
Claudio Fiorino
Dietmar Georg
Günther Hartmann
Tommy Knöös
Antonio Leal
Hans Marijnissen
Ben Mijnheer
Marta Paiusco
Francisco Sánchez-Doblado
Rainer Schmidt
Milan Tomsej
Hans Welleweerd

(Edited by: Mijnheer, Georg)
2008-First edition
ISBN 90-804532-9
©2008 by ESTRO

All rights reserved
No part of this publication may be reproduced,
stored in a retrieval system, or transmitted in any form or by any means,
electronic, mechanical, photocopying, recording or otherwise
without the prior permission of the copyright owners.

ESTRO
Mounierlaan 83/12 – 1200 Brussels (Belgium)
ACKNOWLEDGEMENTS

When writing a booklet about a topic that is under rapid development such as the clinical introduction of an advanced treatment technique like IMRT, there is a chance that at the time of publication many solutions might already be obsolete or replaced by other approaches. We therefore greatly appreciate very much that we got the latest information from the participating centres about the way IMRT verification was implemented, as well as their experience in actual clinical practice. By sharing not only their positive experience with the readers of this booklet, but also by giving examples of errors found by applying their methodology, other centres can profit even more from their know-how. Some of that experience has been collected in a section on pitfalls and potential errors observed in these institutions.

In order to develop and disseminate more uniform guidelines for validation of newly designed IMRT techniques, ESTRO started the QUASIMODO (Quality ASSurance of Intensity MODulated radiation Oncology) network. QUASIMODO was funded by the European Commission, Directorate General Health and Consumer protection – Europe Against Cancer Programme, and was part of the ESQUIRE Project: Education, Science and Quality Assurance in Radiotherapy in Europe, Grant Agreements S.12.3222029 and SPC.2002480. This grant was obtained by a proposal carefully written by Germaine Heeren who was able to formulate the needs from the professionals working in hospitals in various European countries for the safe introduction of IMRT. We very much appreciate her efforts to start this activity and her continuous enthusiasm in following the progress of this project.

Drafting a booklet in which the experience in the field of IMRT verification of a large number of institutions is discussed requires the cooperation of many persons. Besides those who contributed to the contents of this booklet by writing sections on a specific topic, and who are mentioned in the list of authors, we realise that many other persons have provided information that has been discussed in this booklet. We are grateful to all our colleagues who have contributed in one or other way to the contents of this booklet.

A special word of thanks is due to the reviewers of this booklet, Dr Jean Moran from the University of Michigan, Ann Arbor, USA, and Dr Tielo Wiezorek from the Clinic of Radiotherapy, University Hospital Jena, Jena, Germany. Besides critically reading the manuscript, they also gave a large number of suggestions for improving the contents of the booklet. We are grateful for their meticulous work, which made the booklet much more valuable.

Ben Mijnheer - The Netherlands Cancer Institute-Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands
Dietmar Georg - Medical University Vienna-AKH Wien, Vienna, Austria
CONTENTS:

ESTRO BOOKLET NO. 9:
GUIDELINES FOR THE VERIFICATION OF IMRT

AUTHORS
ACKNOWLEDGEMENTS V
CONTENTS VII

1. INTRODUCTION 1
   1.1 Verification of IMRT 1
   1.2 The QUASIMODO experience 4
   1.3 Contents of the report 5

2. VERIFICATION PROCEDURES AND DATA ANALYSIS 7
   2.1 From 1D to 4D verification 7
   2.2 Data analysis 11

3. DOSIMETRY SYSTEMS APPLIED FOR IMRT VERIFICATION 16
   3.1 Ionisation chambers 17
   3.2 Radiographic films, radiochromic films and computed radiography 22
      3.2.1 Radiographic film 22
      3.2.2 Radiochromic film 25
      3.2.3 Computed radiography 29
   3.3 Two-dimensional arrays 30
   3.4 EPIDs 34
   3.5 Gel dosimeters 38

4. VERIFICATION OF IMRT DELIVERY AND TREATMENT PLANNING SYSTEM PERFORMANCE 41
   4.1 Linacs 41
      4.1.1 General aspects of the MLC 41
      4.1.2 Step-and-shoot treatments 45
      4.1.3 Sliding window techniques 46
   4.2 Other types of IMRT delivery systems 47
      4.2.1 Helical tomotherapy machines 48
      4.2.2 CyberKnife systems 52
   4.3 Treatment planning systems 55
5. INDEPENDENT DOSE CALCULATIONS APPLIED FOR IMRT VERIFICATION
   5.1 Monte Carlo calculations 57
   5.2 Other methods 64

6. PATIENT-SPECIFIC QA PROCEDURES AT DIFFERENT CENTRES 68
   6.1 Pre-treatment verification 68
      6.1.1 Ghent University Hospital, Ghent, Belgium 68
      6.1.2 Santa Maria Nuova Hospital, Reggio Emilia, Italy 70
      6.1.3 Tübingen University Hospital, Tübingen, Germany 71
      6.1.4 Seville University Virgen Macarena Hospital, Seville, Spain 73
      6.1.5 University Hospital U.C.L.-St. Luc, Brussels, Belgium 74
      6.1.6 Medical University Vienna-AKH Wien, Vienna, Austria 75
      6.1.7 German Cancer Research Center, Heidelberg, Germany 78
      6.1.8 San Raffaele Hospital, Milan, Italy 80
      6.1.9 Lund University Hospital, Lund, Sweden 81
      6.1.10 University Medical Center Hamburg-Eppendorf, Hamburg, Germany 82
   6.2 In vivo dosimetry 83
      6.2.1 Copenhagen University Hospital, Copenhagen, Denmark 84
      6.2.2 Centre Antoine-Lacassagne, Nice, France 86
      6.2.3 Netherlands Cancer Institute, Amsterdam, The Netherlands 86

7. GUIDELINES 89
   7.1 Accuracy in conventional radiotherapy and IMRT 89
   7.2 Tolerance and action levels of tests for IMRT verification 94
   7.3 Possible pitfalls and potential errors traced by IMRT verification 101
   7.4 Different strategies for patient-specific IMRT verification 103

8. REFERENCES 107

APPENDIX
List of websites of companies selling tools for IMRT verification 127
1. INTRODUCTION

1.1 VERIFICATION OF IMRT

IMRT is currently implemented in a rapidly growing number of centres in Europe. As distinct from the situation in the US, IMRT was until recently applied in Europe in only a relatively small number of, mainly academic, institutions. The reason for this difference in widespread implementation of IMRT was that in Europe IMRT was considered more as an experimental type of treatment technique requiring considerable resources including highly skilled physicists, radiation oncologists, radiation therapy technologists and engineers. Also the special hard- and software necessary for the planning and delivery of IMRT was, at its introduction, still in a developmental stage. Furthermore the routine clinical use of this complex treatment modality required an extensive, time-consuming, acceptance testing, commissioning and quality assurance (QA) programme. In recent years both the hard- and software have become more mature, while at the same time more experience has become available with respect to QA of IMRT. As a consequence many more institutions, including smaller and busy clinics, started with IMRT in Europe, and are now facing the problem of performing a comprehensive QA programme before and during the implementation of IMRT in routine clinical practice.

Since its introduction, many articles, reports and books have discussed issues related to the routine clinical use of IMRT (e.g., the report of the Intensity Modulated Radiation Therapy Collaborative Working Group, 2001, Ezzell et al., 2003, Galvin et al., 2004, Bortfeld et al., 2006). The state of the art of IMRT around 2003, covering numerous technical, physical and clinical aspects, has been discussed in detail during an AAPM Summer School (AAPM, 2003). All these reports emphasize the importance of performing a comprehensive acceptance testing, commissioning and QA programme of IMRT equipment. The need for these types of verification programmes has been demonstrated, for instance, during an independent dose evaluation performed by the Radiological Physics Centre (RPC) of institutions wishing to participate in a Radiation Therapy Oncology Group (RTOG) IMRT protocol (Molineu et al., 2005, Ibbott et al., 2006). An anthropomorphic head-and-neck phantom, having structures where TLDs and radiochromic film could be inserted, was irradiated according to a treatment plan designed by a specific centre. Roughly one third (48/163) of all of the irradiations performed by 128 institutions that irradiated the phantom failed to meet the criteria: 7% in the ratio measured dose to institutions’ stated dose, and 4 mm in the distance-to-agreement in the high-dose region near the organ at risk. These results clearly demonstrate that institutions vary significantly in their ability to deliver dose distributions that agree with their own treatment plans, and that quality assurance tests play a critical role in IMRT planning and delivery. There is, however, not yet consensus to what extent tests, dealing with issues specific for IMRT, should be performed, as discussed, for instance, by Ezzell (2003), Palta et al. (2003), Ahnesjö et al. (2006) and in a forthcoming ICRU report (ICRU, 2008). IMRT
requires verification of a number of parameters, both in the planning and delivery phase, which is not covered by existing protocols. Acceptance testing and commissioning both of an IMRT treatment planning system and of special hardware for delivery of IMRT, is still in most institutions an ad hoc process. This is because new systems are continuously becoming available, while also there exist no clear guidelines and criteria for the accuracy required. Furthermore, the variation in complexity and clinical practice of IMRT in different centres make it unlikely that a single QA programme would fit the needs for all radiotherapy departments. Finally it should be noted that the attitude towards QA of radiotherapy is evolving. Verification of the correct functioning of the separate components required for a radiotherapeutic treatment is no longer sufficient. The combined use, including the connectivity, of the various hard- and software components determines the overall accuracy and reproducibility of a specific treatment, and requires another QA process than applied until recently. Many centres have therefore developed their own QA procedures for IMRT, and only recently some more specific suggestions for tolerance limits and action levels for planning and delivery of IMRT have been provided (e.g., Palta et al., 2003, Stock et al., 2005, McDermott et al., 2007, Sanchez-Doblado et al. 2007).

Commissioning and quality assurance of an IMRT treatment planning system is complex and time consuming, as for any other three-dimensional (3D) planning system, as discussed for instance by Sharpe (2003). Recently several national and international reports, including the IAEA Report TRS 430 (IAEA, 2004), ESTRO Booklet No. 7 (Mijnheer et al., 2004), NCS Report 15 (NCS, 2006) and IAEA TECDOC-1540 (IAEA 2007) provide practical information on the verification of 3D treatment planning systems. These reports allow institutions in a relatively straightforward way to perform tests of their (new) planning system to guarantee the correct functioning of a number of functions relevant for accurate treatment planning purposes. Although the main purpose of the ESTRO booklet, as well as of the other reports, was to provide practical examples for non-IMRT photon beams, some general parts related, for instance, to anatomical description and beam description, are also valid for IMRT. Furthermore the sections of these reports describing dosimetric tests have a number of features that should in any case be verified before using the system for IMRT applications. However, for the additional problems encountered in treatment planning of IMRT, such as the use of small segments and of dynamic delivery systems, additional tests are required to verify the accuracy of dose calculations for those conditions specific for IMRT. Because of the complexity and magnitude of the testing required, both of the specific IMRT issues and the non-IMRT aspects, coordinating these activities with the vendor is extremely important, as discussed for instance in the ESTRO and IAEA documents. Also contacts between institutions having the same system, for instance by means of organised users groups of such a system, may provide valuable information and save a lot of duplication of tests of common software. For instance, extensive testing of algorithms of the system can be avoided in this way, while efforts should concentrate on testing those aspects relevant for the treatment techniques applied in a specific institution.
The International Electrotechnical Commission, IEC, published an International Standard on “Requirements for the safety of radiotherapy treatment planning systems” (IEC 2000). Similar to other IEC documents concerning medical equipment, e.g., for linear accelerators, this International Standard defines a number of requirements to be complied with by vendors of such equipment in order to provide protection against the occurrence of safety hazards to patients. Compliance with these requirements should be checked by testing by the manufacturer and demonstrated to the customer. At this moment there are, however, no general guidelines or tests available for the verification of a TPS for IMRT issues specifically. An independent monitor unit calculation programme is recommended by some groups to verify the dose calculation at one point. More sophisticated programmes, e.g., Monte Carlo-based algorithms, may even calculate the dose in three dimensions to verify the dose calculation of the TPS routinely used for IMRT treatment planning, as will be discussed later in this report. Because dose delivery of IMRT fields is strongly dependent on the accurate position of the leaves, such an independent dose calculation only verifies part of the total chain of the planning, data transfer, and delivery process of IMRT. It is therefore always necessary to verify the leaf positions for IMRT delivery and the regular QA programme of accelerators has to be adapted to the special requirements needed for IMRT applications. For this purpose, special tests are used, as discussed in several publications (e.g., AAPM, 2003, Williams, 2003), which will be summarised in this report. As a result a higher accuracy of leaf position is required and can be obtained than specified according to the manufacturer for “normal” use of the accelerator.

Generally the combined procedure of planning and delivery of IMRT is verified, without performing specific tests of the planning process alone. It is, however, not clear how much effort should be spent in performing tests for IMRT verification. Should IMRT verification be performed for each patient, should all fields be verified or only the total plan, should the dose distribution be verified in one plane, in several planes or the total 3D dose distribution? Is it necessary and practical to perform in vivo dosimetry during IMRT? Currently several groups in Europe are in the process of drafting guidelines for the verification of IMRT. These efforts include working groups of hospital physicists in France (Zefkili et al., 2004), Italy, Spain and the UK, preparing documents written in their national language. The French document has been drafted by physicists of the (GORTEC) group (Tomsej et al., 2005), and gives recommendations for a head and neck IMRT quality assurance protocol. Some of the other reports do not only discuss physics aspects of IMRT but also include staffing requirements (UK) or clinical aspects, for example criteria for patients to be admitted to an IMRT protocol, and margin recipes of volumes of interest (Spain). The QUASIMODO project, which will be discussed briefly in the next section, has collected a lot of experience in IMRT verification in Europe. However, all these activities have not yet resulted in a comprehensive set of practical guidelines for the verification of IMRT. It is the main purpose of this report to provide groups starting with IMRT, as well as groups that have already some experience with IMRT, guidance in the performance of verification tests, as part of an IMRT QA programme, to en-
sure the safe delivery of IMRT to each patient. Attention will be focused on practical issues, for instance by giving examples of tests routinely performed in various European institutions, while theoretical aspects will only briefly be discussed.

It should be noted that a number of IMRT verification methods have been developed at a level that is more accurate than dose calculations performed by commercially available treatment planning systems. For instance in a number of these systems the actual shape of the MLC leaves are not modelled correctly and the transmission through the leaves is not taken into account. Moreover, the further the delivery situation is from 3DCRT, the scenarios that need to be tested continue to increase. These issues should be considered when designing an IMRT verification programme.

1.2 THE QUASIMODO EXPERIENCE

For the verification of a specific IMRT technique in an institution, several approaches can be followed, as will be elucidated in Chapter 2. After performing acceptance testing and commissioning of the various aspects of the IMRT planning and delivery system separately, each IMRT technique should in principle be checked by measuring the 3D dose distribution in a phantom, and comparing its characteristics with those of the prescribed plan. Verifying the dose distribution of beams of an IMRT plan separately is not sufficient because differences between delivery and planning errors might interact in an unpredictable way, thus causing unacceptable errors.

In order to develop and disseminate more uniform guidelines for validation of newly designed IMRT techniques, ESTRO started the QUASIMODO (QUality ASsurance of Intensity MODulated radiation Oncology) network. Fifteen European centres, which just started or were in the process of setting up IMRT, participated in this EC-sponsored project. At the start of the QUASIMODO project it became clear that the strength of a common European project would not be in verifying the individual beams, which should already be part of the QA programme of each institution, but that verification of the composite plan is of utmost importance for the actual patient treatment. Therefore, applying a multi-plane verification using a special phantom, in combination with film and ionisation measurements, approximated 3D verification (Bohsung et al., 2005, Gillis et al., 2005). The results of the QUASIMODO project demonstrated that the dosimetric verification of a complete IMRT treatment is a workable quality assurance tool. Film dosimetry, combined with a suitable data analysis method, allows the user to pinpoint unacceptable deviations between delivery and planning. The approach of a central service centre that prepares/develops all films and performs the numerical comparisons to the respective computed dose distributions, enhances the reliability and validity of such an intercomparison. The agreement between measured and computed dose distributions was better than might have been expected for the wide range of planning and delivery system combinations used by the participating institutions. Systematic differences between planned
and measured dose distributions could be identified, and in some cases the cause revealed. For instance, the most plausible reason for an under-dosage in part of the PTV found in one centre was a wrongly positioned moveable bar in the couch top not taken into account during the dose calculation. Also in another centre the difference between the measured and calculated dose in some parts of the PTV fell outside the criteria due to an accumulation of errors (see Section 3.2.1). The QUASIMODO dose verification trial demonstrated the feasibility of both a multi-centre quality assurance network to evaluate any IMRT planning and delivery system combination and the methodology involved in such a verification study.

Although there are several limitations in the approach adopted by the QUASIMODO group, and improvements are possible, the experience gained with the QUASIMODO project should now be “translated” into guidelines with respect to verification of IMRT techniques for those institutions starting with IMRT. These recommendations should not only describe the methods and equipment required for these tests, but also provide guidelines for the tolerance and/or action levels of observed deviations between planned and delivered dose distributions. In addition, the type of tests to be performed, either simple for each patient or extensive for a few patients after an introduction period, should be indicated, as elucidated for instance by Palta et al. (2003).

1.3. CONTENTS OF THIS REPORT

Ideally the actual dose delivery, in three dimensions, of patient treatments should be verified after performing a comprehensive acceptance testing and commissioning programme of the various phases of the planning and delivery process of IMRT. At this moment in vivo dose verification of IMRT is only employed in a few institutions and pre-treatment verification of IMRT delivery, applying a large variety of phantom-detector combinations, is more often employed clinically. In this report we will first summarise the various approaches for IMRT verification and the methods of data analysis. After briefly comparing the various techniques for dosimetric verification, quality assurance tests for accelerator and MLC performance will be summarised. In separate chapters the use of independent dose calculations and examples of patient-specific QA procedures will then be discussed. These examples are taken from the various centres that participated in the QUASIMODO project, as well as from some other European centres, and show a large variety of IMRT QA tests performed in daily practice in institutions having already some years of clinical experience with IMRT. The report ends with formulating guidelines concerning type and tolerances of tests to be performed for IMRT verification in relation to the required accuracy. In this chapter also a number of possible pitfalls and potential errors encountered by European centres are described, as well as possible actions to solve these problems. In an appendix a list is given of commercial companies selling QA tools for IMRT verification.
It should be noted that this report is not intended to discuss the many details involved in the clinical implementation of IMRT. Various textbooks are available now describing in a comprehensive way the numerous aspects determining the safe introduction of IMRT in the clinic. For instance, in AAPM Medical Physics Monograph No. 29 the state of the art of IMRT in 2003 has been summarised, while more recently the foundations, techniques and clinical applications of IMRT have been reviewed by Bortfeld \textit{et al.} (2006) and by Meyer (2007). The purpose of this document is to provide guidelines for the physical and technical aspects of QA of IMRT. The ultimate decision to accept a plan or to design a “better” plan has to be made in close cooperation with the radiation oncologist, based on patient-specific action levels. These depend on many decisive factors, such as the position and size of the area that failed to pass certain evaluation criteria, and the dose level in the PTV or OAR. In this document we do, however, not refer further to the role of the radiation oncologist in QA of IMRT.
2. VERIFICATION PROCEDURES AND DATA ANALYSIS

2.1 FROM 1D TO 4D VERIFICATION

Verification of dose calculations performed by a treatment planning system is traditionally carried out with a single-point or a few-point detector system, *i.e.* a 1D verification procedure. Such a verification programme typically consists of measuring the dose at specific points or along lines, *e.g.*, depth dose curves or beam profiles, in various types of water or solid slab phantoms, using beam reference data. Numerous publications about the accuracy of dose calculations performed by commercial treatment planning systems exist, and only a few will be briefly discussed here to illustrate the type of information that can be gathered from these verification programmes. In Section 7.1 the results of some inter-institutional dose intercomparisons, using the 1D approach, will be summarised. These intercomparisons were in general more extensive than a verification of the dose calculation algorithm alone.

In an early American study six different treatment planning systems were compared at four different institutions against measurements in about 200 points for each TPS (Masterson *et al.*, 1991). Overall, good agreement between calculation and measurements was found for all algorithms. However, regions in which discrepancies were observed were pointed out, and areas for algorithm improvement were identified. An important attempt to produce a standard set of beam reference data for testing treatment planning systems was performed by AAPM Task Group 23, who developed a set of beam reference data from a 4 MV and 18 MV photon beam (AAPM, 1995). This data set included a number of test cases that could be used for comparison of calculations with the reported measured data after entering the beam reference data into a specific treatment planning system. The data were used to assess the accuracy of dose calculations performed by various commercial treatment planning systems. However, this test package was outdated in a relatively short period of time because photon beam qualities in the range from 6 to 15 MV were not available in this set, while the high-energy photon beam was measured on an obsolete type of accelerator, for instance not having an MLC. More recently Venselaar and Welleweerd (2001) reported about a similar type of study performed in The Netherlands. A coherent set of beam reference data of 6, 10 and 18 MV photon beams was measured on two linear accelerators. These data served as input data in seven commercially available treatment planning systems, which were clinically in use at that time. Next, a test package was measured which included a “missing tissue” geometry and fields with asymmetrical collimator setting, with and without wedge. Absolute dose values predicted by the different treatment planning systems, in which the measured beam reference data were entered, were compared for all test points with the already available results of the measurements. Criteria for acceptability were exceeded by some systems in the case of the irregular field geometry and the missing tissue geometry. Of interest while not reporting the
whole study, the majority of the systems had difficulties with the dose calculation for asymmetrically wedged fields.

Validation of complex systems such as a TPS with an enormous amount of data is a very cumbersome project for which other approaches than point-by-point comparisons should be available. During the introduction of intensity-modulated beams applied for IMRT, the physics community started to perform more extensive verification in 2D (planes) and even in 3D (volumes). In an early publication from the group of the Memorial Sloan Kettering Cancer Center in New York a concept for IMRT verification was introduced (Burman et al., 1997). This approach included:

- Verification of the planned dose distribution by performing an independent dose calculation;
- Comparison of the planned leaf sequence with that recorded in the MLC log files;
- Confirmation of the initial and final positions of the MLC for each field by the record-and-verify system;
- Comparison of the dose distribution measured in a flat phantom with that calculated by the treatment planning computer for the same experimental conditions;
- In vivo dose measurements.

The philosophy behind this new approach was that patient-specific verification was required for IMRT and that each plan should be checked prior to delivery. This was different from the conventional approach where checks are generally performed during the commissioning process of a new TPS or before the implementation of a new technique, using generic or specific geometries in slab or more anthropomorphic phantoms. Previous intercomparisons were often also not restricted to one specific centre but included national and international dose intercomparison projects, as will be discussed more extensively in Section 7.1. These intercomparisons usually applied special phantoms or geometries valid for a treatment technique of a specific patient group. Extrapolation of the uncertainties found during these tests could then be transposed to the planning and delivery of the same treatment technique for individual patients. One of the most important issues in the verification process of IMRT is therefore to make a decision about what efforts should be made with respect to verification of individual IMRT plans compared to the more general QA process applied for 3D conformal techniques. In this report we will discuss various approaches adopted by European centres, both during the initial implementation of IMRT and later after having some years of clinical experience.

A large number of new methods, techniques and detector systems have been designed for patient-specific IMRT verification purposes, where both the dosimetric and the spatial uncertainty of these techniques are determined. Recently De Wagter (2006) and Moran and Xia (2006) presented a review of these techniques, summarised from a European and American
perspective, respectively. Most of these methods still rely on the use of ionisation chambers for assessment of absolute dose values in IMRT fields. Because of the special shape and delivery of IMRT fields, it is not a priori certain that ionisation chamber measurements can be analysed using the procedures outlined in existing dosimetry protocols. In Section 3.1 the results of some studies are presented yielding information on the use of ionisation chambers, and some other detectors, in IMRT fields.

Two-dimensional dose verification methods, such as the use of radiographic film, have gone through a revival as a result of the well-known characteristics of these systems, as well as the ease of the handling process. As a consequence a number of problems and limitations associated with film have been re-experienced again. This revealed process has, however, initiated a number of more profound studies, which have contributed to an improved knowledge of the characteristics of film dosimetry systems for dose verification of IMRT, as will be discussed in Section 3.2. Verification of the dose delivery of separate beams resulted also in the (further) development of a number of a 2D detectors adapted to the special needs for IMRT. Good spatial resolution, fast response and easy analysis of the measured data were a prerequisite for their application as tools for dosimetric verification of individual treatment plans, as elucidated in Section 3.3. The use of electronic portal imaging devices, EPIDs, for dosimetric verification of treatment plans was already proposed in 1990, even before IMRT was discussed (Wong et al., 1990). Dosimetric applications based on electronic portal imaging techniques have been suggested by several groups, both for camera-based, ionisation chamber matrix and amorphous silicon detector systems, as will be discussed in Section 3.4.

Verification in multiple planes, and in some cases in real 3D, is not a common procedure for the verification of 3DCRT and certainly not for individual patients. In fact one could argue why so little attention has been paid in verifying 3DCRT when using more or less advanced techniques. It is the general conviction that performing a thorough commissioning programme of both the planning system and the treatment machine allows the safe introduction of advanced 3DCRT techniques in the clinic. Such a programme should identify the limitations of for instance the dose calculation algorithms of the TPS for specific 3DCRT treatment techniques. In addition to the higher complexity of the dose calculation of the TPS for IMRT compared to 3DCRT, also the accuracy and reproducibility of the delivery of IMRT beams plays an important role. The interaction between these two issues, and the unpredictable effect of uncertainties in both aspects of IMRT on the total dose distribution, is the main reason that with its introduction so much attention is paid to verification in 2D or 3D (see Sections 3.4 and 3.5) of IMRT techniques.

Recently high precision radiotherapy of targets moving during one fraction of a treatment series (4DCRT) is receiving a lot of attention, which is partly related to the development of improved fast imaging techniques. As a consequence 4D treatment planning is developing
(e.g., Keall, 2006) along with 4DIMRT delivery (e.g., Mageras et al., 2006). A number of 4D planning/delivery approaches and tools are currently investigated and verification of new treatment techniques applying the results of these studies is essential (e.g., see Section 4.2.1). Special phantoms have already been described for measuring the dose in moving objects (e.g., Kashani et al., 2007) and respiratory motion phantoms are already commercially available (see the appendix for a list of websites of companies selling tools for IMRT verification). Due to the complex movements of target volumes and OARs, in vivo dose verification of 4DIMRT might even become more important than for IMRT of non-moving targets. At this moment it is, however, not yet clear to which extent the methods presented in this booklet may also be applied for the verification of 4DIMRT.

IMRT delivery can also be performed using compensators, as discussed for instance by Webb (2006). Compensator IMRT techniques require custom-made compensators for each beam, which need special hard-and software for their design, but can nowadays also be ordered at commercial companies. The use of compensator-based IMRT techniques offers certain advantages over MLC-based techniques such as robustness and simpler QA procedures as discussed by Chang et al. (2004) and Webb (2006). Ultimately the resulting 3D dose distribution has to be verified in a similar way as for other IMRT techniques and therefore a number of QA approaches as discussed in this report are also valid for compensator-based IMRT delivery.

IMRT delivery using either the step-and-shoot or sliding window technique may result in considerable longer treatment times than used for 3DCRT treatments. For that reason accelerator companies are now developing arc techniques in order to deliver IMRT in a faster way. Intensity-modulated arc therapy (IMAT) is already applied in a few clinics for some time, mainly for specific tumour sites (see Section 3.2.2). Also during tomotherapy the radiation is delivered while the gantry is rotating. Verification of these rotational IMRT techniques can be performed by applying special phantoms in which radiochromic film is positioned, as elucidated in Section 3.2.2 for an IMAT treatment, or radiographic film, as discussed for helical tomotherapy treatments in Section 4.2.1. Several of the on-line 2-dimensional arrays discussed in Section 3.3 are currently under investigation for their use for pre-treatment verification of arc techniques (see van Esch et al., 2007).

As a final remark it should be noted that it is important, but sometimes difficult, to determine if discrepancies between measured and planned dose distributions are caused by the treatment planning system, the data transfer, the treatment unit, the measurement device or the data analysis. Detailed knowledge of each part of the QA process, including the accuracy of the TPS, MLC position and verification equipment, is required to identify the cause of observed deviations.
2.2 DATA ANALYSIS

The quantitative comparison of dose distributions (e.g., calculated versus measured, or Monte Carlo simulation versus “standard” calculations) has become a key issue in multidimensional dosimetry with the implementation of IMRT. Simple evaluation by superimposing isodose distributions can only highlight or indicate areas of disagreement but does not allow specifying the level of agreement/disagreement in a quantitative way. The most often applied dose evaluation tools comprise a direct comparison of dose differences, a comparison of distance-to-agreement (DTA) between measured and calculated dose distributions, and a combination of these two parameters: the gamma evaluation method. Besides these three commonly applied methods also other dose evaluation tools have been proposed such as the $\chi^2$ confidence interval method (Ma et al., 2003), the normalised agreement test (Childress and Rosen, 2003) and the dose-gradient compensation method (Moran et al., 2005a). These methods have each their own merits but will not be discussed in detail.

Dose differences can be expressed in many ways. Sometimes the absolute value of the dose difference is of interest, but generally the difference is normalised to the dose having a specific value, for instance the prescribed dose, the maximum dose or the dose on the beam axis at the same depth. It should be clear that such a normalisation is not reflecting the local dose difference, which might be a quantity more relevant for organs at risk. In regions of low dose gradients it is sufficient to evaluate dose differences independently of spatial considerations. In regions of high dose gradient, (normalised) dose differences are less meaningful and should be translated into a DTA, as for instance applied in reports on quality assurance of treatment planning systems (e.g., Van Dyk et al., 1993, Venselaar et al., 2001, Mijnheer et al., 2004). These two approaches have to be adopted for the verification of separate intensity-modulated beams or composite (multi-beam) treatment plans where low-dose gradient and high-dose gradient regions can alternate. For that purpose, Low and colleagues proposed the $\gamma$-evaluation method for the quantitative evaluation of two-dimensional dose distributions (Low et al., 1998). This concept combines a dose-difference criterion with a distance-to-agreement criterion for each point of interest. Since its introduction, the $\gamma$-evaluation method has been used for the commissioning of IMRT equipment and patient-specific quality assurance procedures. Refinements of the concept of the gamma evaluation or its application have also been described. For example, Depuydt et al. (2001) applied the $\gamma$-evaluation method for the verification of single IMRT beams with an electronic portal imaging system. They categorised the evaluated points in different filter levels to either reduce the amount of calculation time or to use linear interpolation for suppressing artefacts. Finally, they proposed to reduce the continuous nature of the $\gamma$-value to a pass-fail decision for each point of interest. As a result of the calculation a map of passed or failed points is obtained but the quantitative information, i.e. the numerical $\gamma$-value, is lost. Bakai et al. (2003) revised the $\gamma$-evaluation method by introducing dose-gradient dependent local acceptance thresholds. Low and Dempsey (2003) examined the behaviour of the $\gamma$-distribution in the presence of noise, as present in
Monte Carlo dose calculations, and evaluated the influence of pixel spacing. In order to avoid artefacts in the $\gamma$-calculation in regions with steep dose gradients, the resolution of the dose matrix and the DTA-criteria need to be considered. Based on their analysis Low and Dempsey (2003) recommended a minimum ratio of 1:3 between pixel resolution and DTA criteria.

Besides the correct application of the concept and definition of tolerance and acceptance criteria, the interpretation of a two- or more-dimensional $\gamma$-value matrix is essential. Stock et al. (2005) investigated 10 IMRT hybrid plans verified with film in a polystyrene phantom. Based on the results of these plans with measurements in 3 planes each, they developed a decision filter looking at $\gamma_{\text{mean}}$ values, the average number of pixels with $\gamma > 1$, and the maximum $\gamma$ value expressed as the 1st percentile ($\gamma_{1\%}$). In addition, $\gamma$-area histograms were used for each plane where a comparison between calculated and measured dose distributions was performed. In this way a reduction of the multi-dimensional information concerning the agreement between a reference (e.g., measured) and an evaluated (e.g., calculated) dose distribution seems to be feasible. From $\gamma$-area distributions or histograms statistical data can be calculated to define acceptance criteria for composite IMRT plans or single IMRT beams. Nevertheless, a thorough experimental IMRT-verification needs more than calculating the $\gamma$-distribution. Complementary dosimetric information, such as dose profiles and dose-difference maps, should be considered as well in a quantitative analysis of multi-dimensional dosimetric information. In addition to the calculation of the $\gamma$-index itself, Stock and co-workers looked also at the $\gamma$-angle (see Figure 2.1).

If $D_m$ is the measured dose at co-ordinate $r_m$, $D_c$ the calculated dose at co-ordinate $r_c$, $\Delta D_m$ the dose-difference tolerance criterion and $\Delta d_m$ the distance-to-agreement tolerance criterion, then the gamma value is defined as:

$$
\Gamma(r_m, r_c) = \sqrt{\frac{r^2(r_m, r_c)}{\Delta d_m^2} + \frac{\delta^2(r_m, r_c)}{\Delta D_m^2}}
$$

with $\delta(r_m, r_c) = D_m - D_c$; $r(r_m, r_c) = |r_m - r_c|$

$$
\gamma(r_m) = \min \{ \Gamma(r_m, r_c) \} \forall \{ r_c \}
$$
The $\gamma$-angle can be useful for the interpretation of deviations. It indicates the parameter mostly influencing the $\gamma$-value, i.e. the dose difference or the DTA. The angles of $0^\circ$ are defined on the dose-difference axis. For example, if the $\gamma$-angle is between $\pi/4$ and $\pi/2$ the index is dominated by the DTA criteria. The angle is calculated with the absolute values of dose-difference and distance-difference so that the angle is always between 0 and $\pi/2$. Such information is lost if only the absolute value of gamma is considered. Figure 2.2 illustrates a gamma-value distribution and a gamma-angle distribution for a five-field prostate plan, verified in a polystyrene phantom with EDR2 films.

It is important to recognise that the numerical values for the tolerance or acceptance levels for gamma evaluation are influenced by many factors, for example the IMRT equipment and delivery technique itself or the overall QA policy in a department as will be elucidated in Chapter 7.2. A comprehensive analysis of the verification or measurement procedure should therefore be performed before specifying tolerance and acceptance criteria for quantitative evaluations. Winkler et al. (2005) studied the overall performance of the film dosimetry procedure for IMRT applications and looked at the accuracy of film calibration, influences of field size and depth in phantom on film density as well as phantom positioning in order to define local standards. There are, however, some more general factors that have an impact on the specification of these tolerance levels. Firstly, the complexity of the specific verification procedure, for example, of a composite (multi-beam) treatment plan, will necessitate different limits compared to single beam verification. Secondly, for low dose regions where collimator or leaf transmission dominates dose delivery, the resulting gamma-value distribution depends on the normalisation procedure and whether local or global dose difference criteria
are used. As an alternative solution, points with dose values below a defined threshold value can be excluded from evaluation. Thirdly, the issue of image registration is important because the main intention of IMRT treatment plan verification is to validate the treatment planning system on the one hand, and the quality of beam delivery on the other hand. If, however, the measured dose map is registered with the calculated distribution by the use of markers (e.g., stitches on the film), another source of error is superimposed. Slight shifts in phantom positioning and image registration will influence the spatial conformity.

Figure 2.2 Gamma-value (left) and gamma-angle distribution (right) of a five-field prostate plan with superimposed isodose lines (30%, 50%, 70%, 90% and 95%) calculated by the local TPS (from Stock et al., 2005).

When defining criteria for gamma evaluation methods, or establishing an evaluation filter for the whole IMRT verification process, it is essential to consider also the accuracy in (uniform intensity) 3D conformal radiotherapy. These initial tests may highlight possible systematic errors in the verification process or determine the limitations of the TPS, e.g., uncertainties in beam or penumbra modelling or beam data implementation, mechanical uncertainties of the delivery equipment or dosimetric uncertainties in the measurement techniques. Clinical and technical aspects determining dose accuracy in conventional radiotherapy are discussed more extensively in Section 7.1, while tolerance and action levels for gamma evaluation will be discussed in Section 7.2.

At this moment it is not yet clear to what extent multiple 2D analysis is sufficient or a complete 3D verification is required. For instance, by multiple 2D analyses of 400 pre-treatment verifications van Zijtveld et al. (2007) observed in three cases malfunctioning of one of the leaves, causing a dose difference of about 10%. Due to the combination of several IMRT fields the dose differences reduced to 5% in the reconstructed 3D dose distribution. These
authors concluded therefore that for clinical evaluation of the resulting IMRT dose distributions one cannot rely on the comparison of DVH parameters of the 3D dose distribution only, but that evaluation in single planes is important as well. This issue will be discussed in more detail in Section 7.2.
3. DOSIMETRY SYSTEMS APPLIED FOR IMRT VERIFICATION

A large number of dosimetry systems are available for the verification of IMRT. The choice of a specific system depends on the number of treatment parameters to be verified and the extensiveness of the desired QA process. For this choice the use of a “conceptual pyramid” as shown in Figure 3.1, having four different levels of specificity, was proposed by De Wagter (2004, 2006). Ideally, each time a new or modified IMRT technique is introduced in the clinic, QA starts at the top of the pyramid, *i.e.* by applying a 3D verification technique of the whole treatment planning and delivery process. If unacceptable discrepancies are detected between the 3D dose distribution and the results of the 3D dosimetry verification, the next step is to descend the pyramid to a lower, more specific, level. This can be repeated until the error source is revealed. In this chapter the characteristics of the various dosimetry systems will be discussed, while in Chapter 6 a number of examples of their use in clinical practice will be presented.

![Figure 3.1](image-url) (a) Conceptual pyramid that correlates the various levels of dosimetric QA in IMRT. Like the situation for a real pyramid, each level is based on the stability of the underlying levels. The two lower levels can be part of the periodic QA procedures of equipment used for IMRT planning and delivery. For QA of a new clinical IMRT solution, one may start at the top by applying a 3D dosimetric verification of an entire treatment. One descends the pyramid to the lower levels if the 3D dosimetric verification reveals unacceptable discrepancies with treatment planning. (b) Methodology and tools appropriate for each of the levels. (Courtesy Carlos De Wagter, Ghent University Hospital, Ghent, Belgium, and the Institute of Physics).
3.1 IONISATION CHAMBERS

Absorbed dose determinations using calibrated ionisation chambers in combination with a well-established dosimetry protocol, such as the IAEA protocol (Andreö et al., 2000), are generally assumed to be the gold standard in radiation dosimetry. This assumption is well founded thanks to the knowledge of the physical processes involved in the measurement, the reliability of this type of detector, and last but not least, by the accuracy of the results. Under reference conditions, the estimated combined standard uncertainty in the determination of absorbed dose in high-energy photon beams amounts to about 1.5%. Uncertainties for non-reference conditions, e.g. in the determination of output factors or PDD, are somewhat larger as can be concluded from the dose intercomparisons presented in Tables 7.1 and 7.2.

When ionisation chamber measurements are performed in a typical IMRT field, the departure from reference conditions is more dominant compared to non-IMRT beams. Measurements are affected by the variation in water/air stopping power ratios, $s_{w,\text{air}}$, and in perturbation factors correcting for instance for fluence-averaging effects. Recent studies (Sánchez-Doblado et al., 2003) have shown that calculated stopping power ratios agree for MLC-shaped narrow IMRT beams within 0.3% with the reference $s_{w,\text{air}}$ values for a 6 MV photon beam, i.e. are comparable to their associated uncertainty. The contribution of other factors contributing to the overall uncertainty in the dose derived from an ionisation measurement in an IMRT field depends on the beam quality, the specific type of ionisation chamber and its orientation with respect to the beam axis. For a 6 MV photon beam quality and a micro-ionisation chamber, differences ranging from 0 to 3% have been observed when the ionisation chamber is placed inside an IMRT beam (Capote et al., 2004). However, these differences could reach up to 9% relative to the specified dose if a micro-ionisation chamber is located in the penumbra region, under the MLC or if the segment size is comparable to the size of the ionisation chamber. Similar observations were made by Laub and Wong (2003), Leybovich et al. (2003), and Bouchard and Seuntjens (2004). The latter authors performed accurate Monte Carlo, MC, calculations to determine the correction factor $C_{Q}^{\text{IMRT}}$, which converts a reference field ionisation chamber calibration factor into a calibration factor valid for the IMRT field of interest. The calculation method was experimentally validated using a Farmer-type ionisation chamber and radiochromic film in single step-and-shoot as well as sliding window IMRT fields. Their results showed that for a Farmer-type chamber the correction in a single, realistic sliding window field could be of the order of 10%. Since the magnitude of the dosimetric errors is associated with the fluence perturbation effect, all groups recommend the use of small volume ionisation chambers. However, when using small ionisation chambers large corrections for leakage are sometimes required as shown by Leybovich et al. (2003). These relatively large differences in dose measurements generally occur in very small fields and often compensate each other. For instance, when a micro-ionisation chamber is employed, the uncertainty in the dose measured at a point in the PTV or OAR is for the complete treatment in most cases under 3% relative to the dose specified at that point (Laub and Wong, 2003 and
Sánchez-Doblado et al., 2005a,b). It should be noted that in order to take the volume effect of a finite size ionisation chamber into account a volume dose instead of a point dose calculation has to be made with the treatment planning system.

In order to provide quantitative information about the additional uncertainty introduced in the dose determination, the previous investigations were extended more recently by Sanchez-Doblado et al. (2007) to cover a number of clinical cases and other types of detectors and IMRT techniques. These authors performed dosimetric measurements and Monte Carlo calculations with five different detectors for a number of representative IMRT cases, covering both step-and-shoot and sliding window techniques. In this section the main results of their study, which was partly initiated by the ESTRO-QUASIMODO project, will be summarised while the results for some clinical cases will be elucidated in more detail.

A special IMRT phantom was prepared to obtain a simple and reproducible geometry for the measurements and the MC simulations. The dose was measured at the centre of this cylindrical phantom, which was considered to be the verification point. For this purpose a set of inserts was made in such a way that the reference point of each chamber could be placed in an easy and reproducible manner at the verification point. Three different types of ionisation chambers from the same manufacturer (PTW-Freiburg, Freiburg, Germany) were selected: models 30013 (0.6 cm$^3$ Farmer-type), 31010 (0.125 cm$^3$ Semiflex) and 31014 (0.015 cm$^3$ PinPoint). These chambers differ substantially in their volumes but share the same wall material (PMMA and graphite) and inner electrode composition (aluminium). Also the width and the length of these chambers differ, which are also important dimensions for measurements in IMRT fields. Therefore a possible difference in response of these chambers in IMRT fields should be associated only to a fluence-averaging perturbation effect. In addition two solid-state detectors, a natural diamond and a diode, were employed. Measurements of a step-and-shoot IMRT technique were performed at the German Cancer Research Centre (DKFZ) in Heidelberg, Germany, while measurements of the sliding window IMRT technique were performed at the Arcispedale Santa Maria Nuova (ASMN) of Reggio Emilia, Italy. Based on a large number of clinical cases treated with IMRT both at DKFZ and at ASMN, 13 typical cases were selected for the verification measurements, representing a variety of tumour sites and treatment techniques. The point of measurement was chosen at the isocentre. The variation in dose by introducing small variations in position of the ionisation chamber has also been studied, as elucidated elsewhere (Sanchez-Doblado et al., 2007).

The results of the measurements, given as a percentage difference between the measured dose and the dose calculated by the treatment planning system (TPS), are shown in Figure 3.2 for the step-and-shoot (cases 1-7) and sliding window (cases 8-13) treatments. The “TPS dose” was calculated employing an in-house developed software package (VIRTUOS & KONRAD at DKFZ) or commercial software (Varian, Eclipse v.7.3 at ASMN) for the step-and-shoot and sliding window treatments, respectively. Obviously there was a tendency of an over-esti-
mation of the dose calculation with VIRTUOS (1.5% on average) and of an under-estimation with Eclipse (2.0% on average). This finding, however, was not further analysed, since this study was aiming at the uncertainty in the measurements and not at the quality of the TPS.

The data presented in Figure 3.2 show that for some cases the Farmer-type chamber yielded a considerably different result compared to the other two ionisation chambers. Further investigation showed that these differences could not be attributed to uncertainties in detector position. The question arises how these deviations differ for the various fields applied for a specific treatment. Measurements for two cases have been repeated using the three ionisation chambers, and the results for the separate fields are shown in Figure 3.3. Because all chambers were irradiated using the same plan (file), any effect of accelerator instability, as described for instance by Ezzell and Chungbin (2001), would cancel. Differences were calculated relative to the mean dose of the two smaller ionisation chambers and normalised to total dose. Therefore, the overall result of case #1 differed slightly from that shown in Fig. 3.2. This figure demonstrates that the special behaviour of the Farmer-type chamber can indeed be attributed to particular fields indicating that these fields obviously contain improper conditions for the measurement with the Farmer-type chamber. In their paper Sanchez-Doblado and colleagues gave some examples of positions in the penumbra region where the ionisation response is either higher or lower than predicted by the TPS due to the lack of charged particle equilibrium. When the chamber is fully outside the field no systematic behaviour could be observed. In such cases a check based on the relative position of segments with respect to the ionisation chamber can help to understand the detector response (Sanchez-Doblado et al., 2007).
A careful analysis of the data presented in Figure 3.2 showed a substantial increase in uncertainty for all IMRT cases and all detectors compared to the reference situation, which amounts to 1.5% and 2.0% (1SD) for the step-and-shoot and sliding window technique, respectively. For the sliding window technique, a significant reduction of the standard deviation could be obtained by excluding the solid-state detectors indicating a general over-response of the solid-state detectors. A further reduction for the step-and-shoot cases was obtained by excluding the Farmer chamber results, indicating that the volume effect of the Farmer chamber can significantly contribute to the dispersion, an observation that was also made by other groups (e.g., Laub and Wong, 2003 and Escudé et al., 2006). The overall result was that the additional standard uncertainty in measurements performed with ionisation chambers when verifying an entire IMRT treatment is about 1.0 to 1.5%, provided that appropriate small volume chambers are employed.

The experimental set-up was also simulated with Monte Carlo calculations using detailed modelling of the two accelerator heads (Siemens Primus and Varian 2100C/D), simulated using the BEAMnrc/EGS4nrc code. Absolute dose-to-water in a 10 cm x 10 cm reference field and for the IMRT fields was scored at the centre of the IMRT phantom. For each clinical case the absolute dose delivered to the verification point was evaluated by MC simulation for each IMRT treatment, as well as for the reference field. Assuming the MC-calculated dose is the actual dose delivered by the whole IMRT treatment, instead of the average dose measured with the different chambers, a conversion factor \( c \), was calculated, which related the measured dose to the actual delivered dose. This \( c \)-factor is identical to the factor \( C_{Q}^{\text{IMRT}} \) applied by Bouchard and Seuntjens (2004) during their Monte Carlo calculations. Values of \( c \)-factors calculated by employing the measured chamber dose varied between 0.93 and 1.04. Especially for the Farmer-type chamber, the ratio significantly deviated from one. These findings indicate that for ionisation chambers a type-specific correction factor is required, which is, in

**Figure 3.3** Relative differences between the dose measured with the Farmer-type chamber and the mean dose of the two smaller chambers for two clinical cases irradiated with a step-and-shoot technique (case 1) or a sliding window IMRT technique (case 10)
addition, dependent on the type of modulation of the IMRT field. In practice a large variety of situations occur in which the ionisation chamber is positioned at different positions in or near a segment edge resulting in a compensation of effects as shown by Sanchez-Doblado et al. (2007). As a consequence the average c-factor for all segments will in many cases be close to unity and the use of segment-specific factors is therefore not a very useful approach. Accepting an additional uncertainty in the dose determination in IMRT fields when using ionisation chambers is a more practical solution and therefore recommended. It should be noted that despite this additional uncertainty, the accuracy of dose measurements using ionisation chambers, or other devices, is still at a higher level than most of the dose calculations performed by commercial treatment planning systems. For instance, in some of these systems the shape of the MLC and leaf transmission have been modelled in an approximate way, while in some other systems IMRT dose calculations are based on the intended or expected fluence maps and not on the individual segments that are delivered.

From these studies of the characteristics of different types of detectors to be used for IMRT verification measurements, the following observations can be made:

- There is a tendency for Farmer-type chambers to cause a larger deviation from the actual dose delivered by an IMRT treatment than for smaller chambers. This deviation – clearly correlated to the larger volume - is mainly caused by the lack of lateral electron equilibrium and not so much by the averaging of the dose across the detector volume. The effect is more significant for step-and-shoot than for sliding window IMRT treatments and generally due to a better compensation effect in the penumbra. Consequently ionisation chambers having a small volume are more suitable for IMRT verification than chambers with large volume such as Farmer-type chambers.
- The large discrepancies observed in some of the studied cases suggest that it is useful to perform an additional visual inspection of the position of the chamber in individual fields, preferably in combination with examination of fluence maps, particularly if a large volume chamber is the only choice for performing the verification. In this way it is possible to avoid possible uncertainties arising from positioning an ionisation chamber in regions with a large dose gradient.
- The use of solid-state detectors will introduce additional uncertainties particularly during measurements using the sliding window IMRT technique. This discrepancy may be caused by the energy dependence of their sensitivity, while also a dose rate effect might play a role. In addition, diamond detectors require a pre-irradiation dose. Their use is therefore not recommended for the verification of IMRT.
- By considering only ionisation chambers with small active volume, the additional standard uncertainty in the dose determination due to volume effects is about 1.0 to 1.5%.
3.2 RADIOGRAPHIC FILM, RADIOCHROMIC FILM AND COMPUTED RADIOGRAPHY

3.2.1 Radiographic film

Radiographic films have been employed almost since the discovery of X-rays to measure radiation dose. The use of radiographic films is relatively easy, quick and cheap and therefore very often applied for many applications in radiotherapy. It provides data with a high resolution and a permanent record of the 2D dose distribution in the plane of irradiation. There are, however, many parameters influencing the film irradiation, film processing and data analysis procedure that determine the accuracy of the final result. A number of recommendations made for each of the steps required for film dosimetry has recently been given in the AAPM TG-69 report (Pai et al., 2007). In addition to providing the basic principles and characteristics of film, processors and scanners, some clinical applications of film dosimetry, including its use for IMRT verification, are discussed in that report.

The most characteristic property of an IMRT beam is its intensity distribution in an orthogonal plane, i.e. perpendicular to the beam direction, while the variation of dose with depth is less specific. Therefore, radiographic film seems an ideal 2D dosimeter when oriented perpendicular to the beam axis in a phantom at 5 to 10 cm depth. As optimal depth, planes through the centre of the target volume or of organs at risk can be chosen. For such an orientation, film dosimetry is generally considered to be very accurate in both the high- and low-dose region of the field as observed by many authors (see review of Danciu et al., 2001). For instance, Martens et al. (2002) found that for equivalent field sizes up to 15 cm × 15 cm, deviations between film and ionisation dose determinations remained within 3% for XV2 film (Eastman Kodak, Rochester, NY, USA) for 6 MV and 18 MV beams. However, Yeo et al. (2004), on the other hand, found that both XV2 and EDR2 film (Eastman Kodak, Rochester, NY, USA) exhibit considerable energy dependence in a 6 MV beam. These authors could reduce the over-response in- and outside the penumbra regions from 9% to 3% by using thin lead foils parallel to the film. Part of this inconsistency for film response reported in the literature may be explained by differences in the phantom material and phantom size (Palm and Losasso, 2005).

For the verification of an entire treatment, EDR2 film is preferred over XV2 film because EDR2 film can handle a dose of at least 2 Gy without saturation (Chetty and Charland, 2002), implying that such a generally applied fraction dose can be given to the film. Although radiographic film is widely used as “composite film” dosimeter, its validity is still subject of controversy in the literature and conflicting data have been reported. As a matter of fact, a number of unavoidable problems occur when a multitude of beams contributes to the film response. The essential point is that radiographic film is not water-equivalent, especially due to the silver atoms in the emulsion layers. As a consequence film becomes increasingly sensi-
tive at lower photon/electron energies. This implies that film sensitivity is dependent on both depth and field size of the photon beam, as well as phantom size. Also its use in compensator-based IMRT might therefore introduce an additional uncertainty. An under-response of the order of 5% at 10 cm depth for a 32 mm thick Cerrobend filter in a 15 MV beam was observed by Wiezorek et al. (2005). However, more recently Srivastava and De Wagter (2007) observed a film under-response of only 1.1% for a 30 mm thick block in a 25 MV beam. At lower photon beam energies the effect was even smaller. These deviations are within the experimental uncertainty of routine film dosimetry and indicate that EDR2 film can also be used as a 2D detector for the verification of compensator-based IMRT. Secondly, as discussed by Childress et al. (2002), the varying beam orientation throughout verification of an IMRT treatment, changes the photon spectrum along the film. This introduces a directional differential sensitivity of the film. Indeed, both the XV2 film (Oldham and Webb, 1997, Danciu et al., 2001) and EDR2 film (Gillis et al., 2005) show a higher sensitivity, typically of the order of 4%, when the film is oriented perpendicular rather than parallel to the incoming radiation beam. Robar and Clark (1999), on the other hand, did not find differences larger than 1.5% for a 6 MV beam using Kodak XV2 film for very small field sizes. Another complication with film is the interbatch variability of the emulsion in combination with the variability of the chemical developing process. This effect can be counterbalanced by a rigorous calibration procedure comprising the delivery of a large number of dose values (Gillis et al., 2005) or the adoption of a “universal” relative sensitometric curve that is scaled to an absolute curve by determining the absolute optical density, OD, only at one or two dose values (Bos et al., 2001). Orientation of the film plane during the calibration should be similar to the orientation during the verification process and one should respect a processing time delay to ensure that EDR2 films have reached a stable OD value (Childress and Rosen, 2004).

Figure 3.4 a) Transverse and sagittal view of the polystyrene slab QUASIMODO (CarPet) phantom conceived for the dosimetric verification of IMRT of prostate cancer. Indicated is the planning target volume (in red) surrounding an organ at risk (in yellow). The sagittal view denotes the tilt in the (longitudinal) Z-direction. The seven films are located at Z=60mm, Z=40mm, Z=20mm, Z=0mm, Z=-20mm, Z=-40mm, Z=-60mm. b) Central CT-slice with contours of OAR and PTV and applied co-ordinate system (from Gillis et al., 2005).
In an inter-centre QA intercomparison of IMRT verification, the European QUASIMODO group used a pelvic phantom that contained seven EDR2 films that were oriented parallel to the beam axes as can be seen from Figure 3.4 (Gillis et al., 2005). In this way also the dose in an organ at risk could be assessed. The original intention was to interpret the “composite film” dose values in an absolute way, i.e. without normalisation using an ionisation chamber measurement. However, also due to the earlier-mentioned problems, the QUASIMODO data required a conversion of the “film dose” of each film in order to obtain the actual dose to water. In Figure 3.5 an example of the evaluation of a verification film measurement is shown, revealing a 4-5% systematic over-estimate by the TPS of the delivered dose. This was found to be due to an accumulation of errors: a 1.5% difference in linac calibration on the day of the measurement, and a less optimal kernel in the treatment planning system.

![Figure 3.5](image)

**Figure 3.5** Left: gamma plot at slice $Z = -40$ mm with isodoselines of the computed dose distribution superimposed. Right: Computed and film-measured dose profiles along the dashed line in the left panel. The measured dose shows two peaks at the two holes that served as reference marks (from Gillis et al., 2005).

For IMRT treatments the dosimetric accuracy outside the PTV may be equally important as the accuracy required inside the PTV (Schwarz et al., 2003). However, until now film is seldom used to verify the dose outside the PTV (Childress et al., 2005) and the interpretation of the film response is generally confined to the high dose area (Budgell et al., 2005). One of the underlying reasons is presumably related to the problems described earlier.

From these observations it can be concluded that the use of radiographic film is an accurate method to verify IMRT dose distributions provided an accurate calibration (sensitometric) curve is available, and the film handling and processing is performed in a careful way. Normalisation of the relative dose distribution obtained from film measurements to absolute dose using a (small) ionisation chamber measurement still seems to be necessary to eliminate possible errors introduced during the film measurements.
3.2.2 Radiochromic film

Although silver-halide film is a well-established tool for multi-dimensional dosimetry, it has inherent disadvantageous photon energy dependence in its dose-response behaviour. Consequently, film calibration depends, besides the nominal acceleration potential of high-energy photon beams, on the spectral composition of the radiation field, which in turn changes with field size and depth. Radiochromic film, on the other hand, is known to be almost tissue equivalent and therefore shows little energy and directional dependence. Those radiochromic film types, which have been explored for radiotherapy applications during the last decade, had, however, other disadvantages. The most prominent ones are cumbersome film handling, sensitivity variations across the film, and the low sensitivity to ionising radiation doses typically used in radiation oncology (Niroumand-Rad et al., 1998). This prevented their use in external beam therapy, and their dominant application in radiation oncology was limited to brachytherapy (e.g., Kollaard et al., 2006).

All types of radiochromic films are based on a radiation-sensitive dye (usually a diacetylene monomer), organised into microcrystals and embedded in a gelatin binder. Upon irradiation, a solid-state polymerisation takes place and the film adopts a progressively blue colour. Recently, a new radiochromic film model GafChromic EBT, was designed by the manufacturer to overcome the limitations in using radiochromic film for external beam therapy QA. It offers a high sensitivity compared to that of previous types of radiochromic film, and allows for clinically relevant doses to be accurately determined, i.e. absolute doses can be determined with a low spectral sensitivity, they provide an excellent image resolution, do not require chemical processing (just as other types of radiochromic film), show improved post irradiation colourisation (time between irradiation and readout should be > 2 h), and can be handled and prepared in room light. For these reasons, EBT film seems to be a promising candidate for high-quality dosimetry in IMRT applications (Todorovic et al., 2006). EBT film is supplied in sheets of 8”x10” and recommended by its manufacturer (International Specialty Products, Wayne, NJ, USA) for a dose range of 2–800 cGy. It consists of two active layers (total thickness 34 µm) separated by a surface layer (6 µm) coated onto a polyester base (97 µm at each side). The atomic composition of the film material (42.3% C, 39.7% H, 16.2% O, 1.1% N, 0.3% Li and 0.3% Cl) for the first time includes the moderate atomic number element chlorine (Z=17), raising $Z_{\text{eff}}$ to 6.98 and suggesting that the photoelectric absorption of low-energy photons will be boosted in order to minimise the energy dependence. The changed chemical composition of the active layer gives the unexposed film a lilac hue, which upon irradiation immediately changes to a blue colour (Cheung et al., 2005), as shown in Figure 3.6. After irradiation the maxima in the absorption spectrum, located at 636 nm and 585 nm, did not shift, as observed by Butson et al. (2005).

In order to measure the response to ionising radiation, radiochromic films need to be scanned. For that purpose commercial flatbed document colour scanners have been proposed (e.g.,
Devic et al., 2005, Paelinck et al., 2007 and Stuertewagen et al., 2008). Such scanners are a relatively cheap alternative to helium-neon laser scanners, which operate at a wavelength close to the peak in the absorption spectrum of radiochromic film. Flatbed colour scanners, with a transparency unit, allow scanning films in the transparent mode with a colour depth up to 48 bits in the red-green-blue (RGB) colour mode or 8 to 16 bits in the greyscale mode with a high spatial resolution. The light source in the transparency unit is usually a white fluorescent xenon lamp, which is located behind a diffuser. A charge-coupled device (CCD) that splits the incident light spectrum into three measurable wavelength bands corresponding to the subtractive primary colours red, green, and blue, is used as light detector. Since the absorption spectrum of radiochromic film exhibits a maximum in the red colour region of the visible spectrum, extraction of red colour channel information from a RGB image represents the most frequently used procedure when using a document colour scanner for radiochromic film dosimetry. Warm up effects, if present for the scanner in use, need to be investigated and considered in the dosimetry protocol. After scanning radiochromic films with document scanners, images can be saved in standard file formats, such as TIFF (Tagged Image File Format). Such files can then be imported into standard commercial software for film dosimetry or further processed in other research software platforms, e.g. MATLAB (The MathWorks, Inc., Natick, MA, USA). Using such software tools, dosimetric comparisons of line dose values (profiles) or gamma evaluation can be performed (see Section 6.1.10 for an example of the clinical application of radiochromic film).

The scanner noise in the final results can be minimised and thus the accuracy of radiochromic film dosimetry can be improved if pixel values, measured in regions of interest, are averaged for each colour channel for multiple successive scans of the same film (Devic et al., 2005). Other practical aspects related to the overall dosimetric accuracy are the homogeneity of the scanner’s response or the film positioning accuracy in the scanner (Fuss et al., 2007, Paelinck et al., 2007 and Stuertewagen et al., 2008). The latter influence can be minimised by using a film template for positioning. Film positioning is crucial because radiochromic EBT films are known to show polarisation effects and therefore calibration curves obtained in landscape or portrait orientation might differ. It is therefore imperative to perform film calibration and film dosimetry with the same film orientation in the scanner, irrespective of the film-reading device.

At Ghent University Hospital, radiochromic film dosimetry is applied for pre-treatment patient specific QA when new IMRT class solutions are clinically implemented, or in case of complex “special” treatments for individual patients. The approach is to deliver the clinical treatment plan to an anthropomorphic phantom loaded with one or more Gafchromic EBT film sheets and to compare the measured dose distribution to the one that is computed for the phantom. To illustrate the procedure, a pre-treatment dosimetric verification is described that was applied to a 5-arc IMAT treatment of an elongated tumour adjacent to the thorax wall. The planning strategy was to spare lung tissue as much as possible. The mixed photon beam
quality, 6 MV (2 arcs) and 18 MV (3 arcs), required an energy independent dosimeter in the MeV range, for which radiochromic film is well suited. First, the treatment plan was transferred to the Pinnacle treatment planning system and was delivered to the QUASIMODO CarPet phantom (see Figure 3.4).

The 3D dose distribution was computed using the convolution/superposition algorithm. Second, the CarPet phantom was loaded with one EBT film sheet in the isocentric transverse plane and the treatment plan was delivered to the phantom. The dose distribution was situated close to the phantom edge. Therefore the film was precisely pre-cut using a computer controlled milling machine to avoid attenuation effects by film parts exterior to the phantom. This technique trims the film very precisely according to the local phantom contours and avoids damage on the film edge. Figure 3.6 displays the CarPet phantom with the EBT film after the delivery of the IMAT treatment. Subsequently, computed and measured dose distributions were compared using the gamma evaluation method. Data analysis was performed with in-house software developed in a MATLAB environment. Figures 3.7.a and b represent the computed and measured dose distribution in the isocentric plane, respectively, while Figure 3.7.d shows their difference. Figure 3.7.c displays the gamma values with the isodose lines of the computed dose distribution superimposed. A dose difference criterion of 3% and a distance-to-agreement criterion of 3 mm were applied. The verification was repeated and a similar gamma plot was found the second time, demonstrating the reproducibility of the radiochromic film dosimetry procedure and the systematic nature of the cause of the deviations.
Based on the gamma plot of Figure 3.7.c, the treatment plan was accepted. Repeating the experiment in a parallel transverse plane 2-mm offset from the isocentre basically resulted in a similar gamma plot. This finding contradicted the hypothesis that the abutment of the two central leaf pairs affected the isocentric transverse plane. Possible further explanations such as the irregular start-up or stop-down transients in gantry rotation as well as the attenuation caused by the tabletop also could not explain the observed differences. More detailed measurements of each arc separately demonstrated an error in the output factor calculation of a specific release of the TPS.

In summary, each film dosimetry system consists of three main components that determine the overall accuracy: the type of film, the scanning device, and the film scanning / evaluation protocol. As an example of a protocol for radiochromic film dosimetry, based on EBT type films and a commercial flatbed scanner (type EPSON Expression 1680 Pro), the one
established in collaboration between the Department of Radiotherapy, Medical University Vienna, and the Department of Radiotherapy and Nuclear Medicine, Ghent University is given in Table 3.1.

Table 3.1 Working protocol for EBT radiochromic film dosimetry using a flatbed scanner (from Stuertewagen et al., 2008).

<table>
<thead>
<tr>
<th>EPSON scanner protocol</th>
<th>EBT Gafchromic film protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use a positioning frame to position the films on the same place</td>
<td>Use gloves to handle the films</td>
</tr>
<tr>
<td>Remove the positioning frame during scanning</td>
<td>Use tight-light envelopes for storage</td>
</tr>
<tr>
<td>Perform at minimum 5 successive scans before real measurements</td>
<td>Cut film pieces at minimum one day prior to irradiation</td>
</tr>
<tr>
<td>Turn the scanner off between the measurements</td>
<td>Use the films in portrait orientation</td>
</tr>
<tr>
<td>Use the same specifications in the EPSON software: professional mode, transparent document type, set 48-bit. colour correction off; select 150 dpi resolution</td>
<td>Scan the films before and after irradiation and use the net optical density for dosimetric evaluation</td>
</tr>
<tr>
<td></td>
<td>After irradiation wait at least 4 hours to scan the films</td>
</tr>
<tr>
<td></td>
<td>Use/select the red colour channel</td>
</tr>
<tr>
<td></td>
<td>Use MatLab software to obtain and process the measured pixel-values; including a 2D correction for scanner inhomogeneities (due to variations in light scattering).</td>
</tr>
</tbody>
</table>

3.2.3 Computed radiography

Following the transition in hospitals to digital replacement of radiographic film, computed radiographic (CR) plates have been considered for 2D dosimetry. The active layer of a CR plate is a photostimulable phosphor (BaSrFBr:Eu$_{2}^{2+}$) approximately 300 μm thick. These phosphor storage foils have a typical size of 24 cm x 30 cm and a thickness of 0.9 mm, and have to be inserted into a light tight envelope, as room light exposure results in signal decay. Olch (2005) studied Agfa CR plates (model MD 10) in combination with a Kodak 2000RT CR scanning system to dose values up to 328 cGy. The dose-response curve is semi-logarithmic up to 150 cGy, beyond which it flattens. Also other characteristics such as reproducibility and spatial uniformity of the response, signal decay with time, perpendicular versus parallel irradiation, and field size dependence of the response were investigated by Olch. Due to the high-Z constituents of the active layer of the CR plate, current phosphor storage
media feature a dose-response enhancement of lower energy scatter dose for larger field sizes and at larger depths, similar to radiographic film. From his study Olch concluded that with careful attention to methodology, the system accurately performs for both relative and absolute dosimetry for single open beam and IMRT beam QA tests with field sizes less than 15 cm x 15 cm. For composite IMRT QA tests, relative dosimetry was performed with only slightly reduced accuracy compared to radiographic film. CR dosimetry is therefore certainly an option for replacement of film dosimetry for departments that already moved to (portal) imaging using storage phosphor imaging plates (Geyer et al., 2007).

3.3 TWO-DIMENSIONAL ARRAYS

Simultaneous dosimetric measurements in more than one point have become an important need for quality assurance in modern radiotherapy applying time variable fluence patterns, e.g. for dynamic wedges or IMRT (e.g., Watts, 1998, Amerio et al., 2004, Stasi et al., 2005). Such measurements are traditionally performed with radiographic films as a two-dimensional detector. However, their application is not straightforward due to many factors of influence on the optical density, such as energy and spectral composition, depth, field size, orientation, and processing conditions as elucidated in Section 3.2.1. Additionally, the increasing number of IMRT patients suggests the use of faster and more efficient dosimetric tools. Finally, many hospitals are aiming towards a so-called “digital hospital”, where film-processing machines for traditional silver-halide films will not be available or easily accessible in the near future. Radiochromic films, which are self-developing, represent an alternative to radiographic films but their use is still limited because they were until recently rather expensive. The advantages and disadvantages of film as a dosimeter have been described in more detail in the former sections.

The major benefit of dosimeter arrays are their simple handling by connecting them to a PC and the availability of on-line information. Although one-dimensional arrays have been investigated for their use in IMRT (Martens et al., 2001), two-dimensional (2D) arrays are more practical as they allow the verification of a planar fluence or dose distribution. During the last years several systems became commercially available for 2D dosimetry. The most commonly utilised dosimetric principles are ionisation in air or ionisation in semiconductor material, but other principles such as scintillation have been applied as well. Advantages of ionisation chambers are the simple calibration, practically no dead time, which allows real time measurement, and no (significant) effect of radiation damage. In general, dosimetric properties are governed by the physics principle of the detector. The most important ones are dose linearity, energy dependence, directional dependence and dose per pulse response. The latter characteristic is especially important for IMRT, as the beam fluence changes in an IMRT beam can be hundredfold (Jursinic and Nelms, et al., 2003).
The various commercially available 2D arrays show differences in the number of detectors, detector spacing, detector shape, effective point of measurement, water-equivalent build-up layer, backscatter layer, and maximum field size covered (e.g., Letourneau et al., 2004, Spezi et al., 2005, Wiezorek et al., 2005 and Poppe et al., 2006a,b). Electronic portal imaging devices are another option for two-dimensional dosimetry. These devices and their application for portal dosimetry are discussed in the next section. Besides the difference in the underlying physical principle for dosimetry, properties governed by the read-out electronics can play a role for specific investigations. For example, the minimum read-out time of electronic chips might not support investigations of start-up characteristics of the linac, which is important for step-and-shoot IMRT, or to verify leaf or jaw movements at various speed, which is important for dynamic IMRT delivery. Furthermore, as for any other dosimeter, practical issues such as warming up effects or time stability should be carefully considered prior to application for precision measurements.

The 2-dimensional array systems based on diodes or pixel ionisation chambers have detectors located at fixed positions, with grid spacing between the sensitive volumes ranging from 7 to 10 mm. With respect to spatial resolution, the Nyquist theorem can be applied to relate the detector spacing and spatial frequencies in the dose distribution. Most of these systems can be used for absolute dose measurement after appropriate individual calibration procedures to correct for response variations across the array. Detector spacing and thus the resolution have been discussed in various articles (e.g., Stasi et al., 2005 and Herzen et al., 2007). This characteristic has some practical implication, for example when verifying leaf calibration. If the detector spacing does not correspond to the leaf width, dedicated gantry holder devices can be used to “adapt” the projected MLC width to the pixel spacing. An alternative solution might be to adapt the SSD as discussed in Section 4.1.1. Furthermore, 2D arrays mounted on the gantry enable IMRT verification at gantry angles identical to the ones applied in treatment plans. For such procedures detector misalignments and influences of gravity need to be considered carefully and corrected for if present.

The commercial 2D arrays for dosimetric purposes come with their inherent evaluation software. It is generally possible to import calculated dose distributions from a planning system and to perform 1D dosimetric analysis of profiles or a 2D gamma evaluation using data of the whole array. The major limitation of 2D arrays is their limited number of detectors, which impairs measurements in high dose gradient regions and in small fields. The usefulness of a gamma evaluation, based on dosimetric information with a limited spatial resolution is therefore questionable.

Figure 3.8 shows dose profiles of an IMRT verification procedure (for the same calculated intensity profile) using films and different commercial 2D detector arrays, while Figure 3.9 shows an example of both a gamma evaluation and a comparison along a profile. Obviously the limited spatial resolution of the 2D array influences the effectiveness of the verification at
some points. For that reason it is recommended to combine the results of multiple measurements in which the array has been replaced over a small distance. Several other examples of using 2D arrays in clinical practice are given in Chapter 6.

The individual detector response has been investigated for pixel ionisation chamber systems. More specifically, narrow slit beams (in the mm range) have been scanned over single chambers to determine experimental response functions $g(x,y)$ of such a single detector. For dosimetric IMRT verification the exported dosimetric information from the treatment planning system was then convolved with the response function prior to comparison with the measured dose distribution using a 2D array (Poppe et al., 2006a, Herzen et al., 2007). The necessity of such an approach was explained by the construction of 2D arrays with a limited resolution and photon interactions with the (non-water equivalent) material in between chamber elements, where electrons emerge that can be laterally scattered and transported to the air-filled sensitive volumes.

**Figure 3.8** Example of an IMRT verification (for the same intensity profile) performed with different commercial 2D detector arrays. All intensity profiles marked as “calculated” refer to IM profiles obtained with the TPS. Measurements were made at 10cm water equivalent depth with radiochromic film (EDR2, left upper), a diode array (Mapcheck, right upper), a scintillation detector (I’mRT, left lower) and an ionisation chamber array (Seven29, right, lower). The 10 cm water equivalent depth included the inherent build-up of the 2D detector arrays. For comparison EDR2 film measurements are shown as well (from Wiezorek et al., 2005).
Recently transmission-type radiation detectors have been developed that can be positioned on the radiation entrance side of the patient (Poppe et al., 2006b). These detectors are multi-wire or multi-strip ionisation chambers connected to a multi-channel electrometer. They are designed to be placed in dedicated holders or in standard accessory holders of the linear accelerator, e.g. at the position of shielding blocks. As a consequence the spatial resolution depends on the mounting distance. Because of their negligible attenuation, transmission-type 2D detectors can be permanently installed on accelerators primarily used for IMRT. They enable on-line monitoring of beam characteristics or leaf settings with and without the patient in place.

In summary, it is generally accepted now that there is potential to optimise and “modernise” IMRT verification procedures by using electronic real-time two-dimensional detector systems. However, when using such detectors the characteristics of a certain device need to be taken into account, including its specific influences on the overall QA procedures and dosimetry logistics (Wiezorek et al., 2005). Moreover, besides the advantage of offering on-line information, 2D detectors have the potential to increase the overall efficiency for IMRT QA. In addition, these tools can also be used for QA of linear accelerators used for conventional treatments, such as measurement of leaf position, output constancy, beam symmetry and field flatness. 2D detector arrays have also been used for IMRT verification of helical tomotherapy as discussed in Section 4.2. The development of 2D arrays for IMRT verification is currently a dynamic field of research. It can therefore be expected that more information about these systems becomes available in the near future. Results of pre-clinical testing and other information about commercial releases are generally available on the homepage of manufacturers of dosimetric equipment (see Appendix).
Figure 3.9 Verification of an IMRT treatment using a 2D detector array. Top left: measured isodose lines; top right: isodose lines calculated by the TPS. Bottom left: gamma evaluation of the two dose distributions; bottom right: beam profiles along the horizontal green line. At some points differences between the measured dose and the dose calculated by the TPS can be observed due to the finite spatial resolution of the detector array.

3.4 EPIDS

On-line electronic portal imaging devices (EPIDs) have been developed for acquiring megavoltage images during patient treatment (Antonuk, 2002). Megavoltage images, obtained in digital format with such a device, are then used for further analysis, mainly for determining set-up errors. The image information can, however, also be related to the dose delivered to the EPID, yielding dose information in a plane instead of in one or a few points. EPIDs can serve for several purposes during the verification process of IMRT and are used: 1) to verify the leaf position either during static (step-and-shoot) or dynamic MLC (sliding window) techniques; 2) to check the correct transfer of the leaf sequencing file to the treatment machine, and 3) to measure the combined mechanical and dosimetric performance of the treatment unit. The first application will be discussed in Section 4.1 while in this section the latter two aspects will be elucidated.

In the past most studies were performed for three commercial systems: the fluoroscopic, CCD camera-based Philips SRI-100 system, the Plumbicon tube camera-based TheraView (Cablon)
imager and the Varian Portal Vision liquid-filled matrix ionisation chamber system. More recently the use of flat panel imagers based on amorphous silicon, a-Si, are becoming more popular for their use as 2D dosimeters. Before using an EPID for dosimetry purposes, a number of basic dosimetric characteristics have to be determined, such as the dose-response curve, reproducibility of the signal, temperature, dose rate and dose per pulse dependence, and response variation with gantry rotation angle. The results of these measurements show that video-based systems are fast and have a good linear response (Pasma et al., 1999). These systems need, however, a relatively large correction for light scatter in the detector, which is position dependent. The matrix ionisation chamber system has a non-linear response and is relatively slow; i.e., it needs more MUs for the same signal compared with other systems (Boellaard et al., 1998). Amorphous silicon systems have an almost linear response and are fast, as reported by several groups (McCurdy et al., 2001, Grein et al., 2002, Greer and Popescu, 2003, McDermott et al., 2004, 2006a, Van Esch et al., 2004, Moran et al., 2005b, Winkler et al., 2005). Most new accelerators are nowadays equipped with an a-Si EPID and it can therefore be expected that the use of these devices for IMRT verification will increase in the future.

There are two problems related to the use of a-Si EPIDs for dosimetry purposes. The first is how to handle “ghosting”, i.e., the additional signal after the irradiation has been stopped (Siewerdsen and Jaffray, 1999, McDermott et al., 2004, 2006a, Winkler et al., 2005). This effect is not of great concern if an a-Si EPID is used for the verification of fields having a large number of monitor units, because ghosting can be taken into account using a semi-empirical correction procedure (McDermott et al., 2004). It is, however, of importance during the verification of IMRT where individual pixels are irradiated in an irregular manner and therefore difficult to correct for missing signal. The second problem with a-Si EPIDs is the energy dependence of their response. Due to the presence of high-Z materials a-Si EPIDs exhibit an over-response for low energy photons. As a result the sensitivity of an a-Si EPID is field-size and depth dependent, if positioned in a phantom, similar to film dosimetry (e.g., Nicolini et al., 2006). In addition there is an off-axis sensitivity variation (Greer, 2005). By adding a 2 to 3 mm thick copper plate to an a-Si EPID its energy dependence can be reduced (McDermott et al., 2004). Various methods have been developed to take this energy variation of the sensitivity into account. A first approximation is to use the so-called flat- or flood-field calibration of the pixels of the raw image. By using measured data or Monte Carlo calculations a more accurate description of the field size dependence of the sensitivity can be obtained. Empirical correction models (Greer 2005, Nicolini et al., 2006) and convolution models applying kernels describing the X-ray and photon scattering in the EPID (McCurdy et al., 2001, Chang and Ling, 2003, Warkentin et al., 2003, Wendling et al., 2006) have been developed for converting EPID images into dose distributions at the position of the EPID.

Several approaches have been described for the use of EPIDs for pre-treatment verification of IMRT delivery. Pasma et al. (1999) acquired EPID images using a CCD camera-based
fluoroscopic EPID for all beams produced with dynamic IMRT delivery. These images were converted into two-dimensional dose distributions and compared with dose distributions calculated with a treatment planning system at the position of the EPID. A similar forward procedure for pre-treatment verification first for a liquid-filled matrix ionisation chamber EPID (see Figure 3.10) and later for an a-Si EPID, has been explored by Van Esch and colleagues (2001, 2004). The accuracy of their method was within 2% with respect to film and ionisation chamber measurements but became less accurate for higher leaf speeds, particularly in areas of steep dose gradients. Such a forward approach using a-Si EPIDs has also been applied by other groups (e.g., McCurdy et al., 2001, Grein et al., 2002, Chang and Ling, 2003, Greer and Popescu, 2003).

Figure 3.10 Comparison of a dose profile extracted from a liquid-filled type of EPID relative to: (a) a film scan (open symbols); and (b) ionisation chamber measurements in a water phantom (closed symbols). All measurements have been normalised to the point on the beam axis (from Van Esch et al., 2001).

More recently various groups have developed methods to translate EPID images into 2D primary fluence maps, which are then used as input in a TPS to recalculate 3D dose distributions using CT data of a phantom or patient (e.g., Renner et al., 2003, Steciw et al., 2005, van Elmpt et al., 2006 and van Zijtveld et al., 2007). Generally these approaches are able to reconstruct the 3D dose distribution in phantoms with a high accuracy. For instance, van Elmpt and co-workers compared the reconstructed dose for an IMRT plan with the dose distribution obtained with film and ionisation chamber measurements. When using a gamma evaluation with a 3%/3 mm criterion, 99% of the pixels inside the irradiated field had a gamma value smaller than one, whereas the absolute dose at the isocentre agreed within 1% with the dose measured with an ionisation chamber. By calibrating and performing measurements with their EPID at 105 cm SSD, Nicolini et al. (2006) were able to compare directly 2D dose measurements at 3.8 and 1.5 cm depth in a phantom with those derived from a TPS. At The Netherlands Cancer Institute the EPID response is converted to dose inside a phantom or patient using a back-projection algorithm (Wendling et al., 2006). The latter authors compared the results of their dose reconstruction, having a resolution of 1 mm at the isocentre, with measurements in
phantoms using various types of dosimeters. Generally the agreement is very good (see Figure 3.11) indicating that this approach is able to accurately predict the dose in a plane in a phantom and therefore can replace film dosimetry for 2D pre-treatment verification of IMRT. Because the model applies the separate images acquired by the EPID, the maximum number of segments that can be verified is in practice not limited because it is only determined by the (high) image acquisition speed. In Chapter 6 an example is given of the use of this method for in vivo dose verification of IMRT treatments.

![Figure 3.11](image)

**Figure 3.11** Comparison of EPID and film dose distributions inside a phantom for pre-treatment verification of an IMRT field consisting of eight segments using an 18 MV photon beam. The 20 cm thick polystyrene slab phantom was located at an SSD of 90 cm. The EPID dose was reconstructed at a depth of 10 cm. The film measurement was done at the same depth simultaneously with the EPID measurement. (a) Two-dimensional dose distribution derived with the EPID, isodose lines are shown. The vertical line in the dose distribution indicates the position of the y-profiles shown in panel (b), EPID dose as solid line, film dose as dashed line. (c) Two-dimensional γ-distribution of EPID versus film dose. A dose-difference criterion of 2% of the maximum dose and a distance-to-agreement criterion of 2 mm were used (from Wendling et al., 2006).

Another approach translating the EPID dose to a patient dose has been explored by Partridge et al., (2002) and by Pouliot and co-workers (Morin et al., 2006). These groups developed a mega-voltage cone-beam, MVCB, CT imaging system based on the use of an amorphous silicon type of EPID. Although MV CBCT is primarily designed to account for the patient anatomy at treatment time, the images of a calibrated EPID can also be used to obtain an estimate of the actual patient dose. The energy fluence measured with the EPID and the MV CBCT of the patient can be used together to reconstruct the 3D dose distribution delivered to a patient. In this way the effect of both anatomical changes and linac delivery or other types of errors can be assessed. Both groups described a proof-of-principle using a phantom, but the clinical use of the system for in vivo IMRT verification is still under development.
It should be noted that software for portal dosimetry purposes of commercial EPIDs is still only available at a limited extent; a situation that hopefully will change in the near future.

3.5 GEL DOSIMETERS

Most dosimetry systems discussed until now allow verification of IMRT dose distributions in a point (e.g., ionisation chambers), along a line (e.g., linear detector arrays) or in a plane (e.g., film and two-dimensional arrays). EPIDs can also be used for 3D verification as discussed in the former section. A more direct way of verifying 3D dose distributions is by using gel dosimetry, in which a hydrogen containing gel records the 3D dose distribution, which is then read out by means of magnetic resonance imaging (MRI) (e.g., De Deene et al., 1998, Gustavsson et al., 2003) or optical computed tomography (e.g., Oldham et al., 2001). Currently the most viable gel dosimeter consists of gelatine that is doped with monomers that polymerise by absorption of ionising radiation (Maryanski et al., 1996). Magnetic resonance imaging (MRI) allows the quantitative acquisition of the dose distribution according to the relationship between $R_2$ (the reciprocal of the $T_2$ relaxation time) and absorbed dose, which has to be obtained for each gel production. This calibration includes MRI of a number of test tubes containing gel exposed to different dose levels typically between 0 and 8 Gy. A vacuum technique can be used to model the gel cast after a specific patient or anthropomorphic phantom. The resulting gel phantom is irradiated completely according to the treatment plan except for the absolute dose: in order to fully exploit the dynamic range of about 8 Gy, the number of MUs are scaled up, or the clinical treatment plan is delivered a number of times.

Gel is potentially an ideal dosimeter. A gel dosimeter is able to simultaneously integrate the time-dependent 3D spatial distribution of dose delivery during IMRT. The gel dosimeter is water-equivalent and can itself constitute the anthropomorphic phantom, making it free from dose perturbation effects. For relative dosimetry, an accuracy of 3% (1SD) for an in-slice spatial resolution of 1.5 mm and a slice thickness of 5 mm has been reported by the Ghent University research group (De Deene et al., 1998). Although a real patient case can serve as an ideal starting condition, gel dosimetry is too complex and laborious for patient-specific QA. In that view, gel dosimetry is to be considered as a pre-clinical rather than a pre-treatment QA tool to be used during the commissioning phase of a TPS or for each site-specific IMRT technique (class solution) introduced in the clinic (De Deene et al., 2000, Gustavsson et al., 2003).

Duthoy and colleagues (2003) used gel dosimetry for a dosimetric verification of whole abdomenopelvic intensity-modulated arc therapy (IMAT). IMAT is probably the most demanding IMRT method in terms of equipment performance (Williams, 2003). One of the challenges for using gel dosimetry for this situation was the large gel volume that was incorporated in the hybrid phantom displayed in Fig. 3.12.a (Vergote et al., 2004). The 3D capabilities of
gel dosimetry allowed the construction of gel-measured dose–volume histograms (DVHs), which are compared to the computed (planned) DVHs in Fig. 3.12.b. These DVHs were derived from an interpolation separately applied within the computation grid (5-mm cubic voxels) and the measurement grid (1.25-mm transverse and 5-mm longitudinal voxel size). The DVHs of liver and right kidney reveal discrepancies between gel measurements and computations, which are probably due to the discretisation of the gantry angles during the calculation of the arc movement. Gel dosimetry nevertheless confirmed that all the clinical planning objectives were satisfied.

At the moment, gel dosimetry is still in a research phase and it is not widely accepted yet as a mature dosimetry method for pre-treatment verification. Issues that impede the dissemination of gel dosimetry are:

- Labour intensive and costly production of the polymer gel. Especially the need to expel all dissolved oxygen from the gel, polymer gel fabrication requires specific expertise and specialised laboratory equipment. This might change with the use of anti-oxidants to chemically bind the free oxygen as proposed by Fong et al. (2001). These normoxic polymer gels can be produced within an hour, without the need for expensive equipment.
- Limited access to MRI scanners and the implementation of quantitative MRI protocols. To this end, the Ghent University research group has started the distribution of a quality control phantom and an imaging protocol.

**Figure 3.12** Hybrid gel phantom, based on the Alderson-Rando phantom, used for the 3D-dose verification of an IMAT treatment for whole abdomino-pelvic irradiation. The Barex® cast contains 9-liter of polymer gel. The seven reference markers, which are attached to the phantom on the laser lines, contain CT- and MRI-contrast agents. (b) Resulting dose–volume histograms from the gel dosimetry system (dashed lines) compared to the computations (solid lines) in terms of clinical dose after 22 fractions. In order to have maximum precision in gel dosimetry, the number of MUs of the treatment plan had to be scaled up in order to obtain 7.5 Gy as mean dose in the PTV. (from Duthoy et al., 2003).
Due to these limitations of the use of MRI scanners, optical CT scanning has been proposed as an alternative (e.g., Oldham et al., 2001). This method has also its shortcomings such as the introduction of scatter and edge artefacts. For those and other reasons research is still going on to design a convenient and accurate 3D dosimetry system that can overcome these limitations. More recently a radiochromic dosimeter in combination with an optical CT scanning system has been proposed as a high-resolution 3D dosimetry system (Guo et al., 2006). Its use in IMRT has, however, still to be proven.
4. VERIFICATION OF IMRT DELIVERY AND TREATMENT PLANNING SYSTEM PERFORMANCE

4.1 LINACS

Before starting clinically with IMRT one has to verify the integrity of radiation output and mechanical aspects of a treatment machine by use of special QA measures as outlined in Chapter 1. In this section some tests concerning accelerator and MLC performance will be discussed. After the acceptance testing and commissioning of the TPS, sequencer and delivery system, the next step is to verify IMRT treatment plans by dosimetric measurements for each new treatment technique and site (e.g., prostate, head-and-neck, breast, lung) using the “conceptual pyramid” approach as elucidated in Chapter 3. Next, pre-treatment verification should be performed at least in one plane for a number of IMRT plans. Such a procedure might be reduced or even stopped after getting sufficient confidence in the actual beam delivery. Whatever method of patient-specific QA is chosen, each approach should be accompanied by an intensive QA programme of IMRT-specific machine equipment, including a check of the transfer of data from the treatment planning system to the treatment machine. Deviations between actual and prescribed leaf position may be caused by various reasons, such as changes in optical system and leaf motor stability, as discussed for instance by Vieira et al. (2006). Relatively small deviations can already be clinically significant as shown by various authors (e.g., Sastre-Padro et al., 2007). For step-and-shoot the output and beam profile have also to be checked for small number of MUs, while for the sliding window technique not only the leaf position but also the leaf speed has to be verified regularly. In this chapter a number of tests for verifying the MLC performance will be presented.

4.1.1 General aspects of the MLC

Several reports have been published describing a variety of tests of MLC performance as well as constancy of MLC parameters (e.g., Chui et al., 1996, LoSasso et al., 1998, LoSasso 2003, Galvin et al., 2004). Various approaches can be applied for verifying the correct leaf position. MLC leaf position deviations of about 0.2 mm can be measured using match-line uniformity (Galvin et al., 2004). The so-called “garden fence test” or strip test as introduced by Chui and colleagues for sliding window IMRT delivery, can also be used for the step-and-shoot mode (Sastre-Padro et al., 2004). If the test is performed in a relative method it verifies whether all leaves are having the same position relative to the line through the isocentre perpendicular to the leaf movement. The test can be performed using films or an EPID, and is able to detect calibration errors of +/- 0.5 mm just by visual observation. Mathematical analysis of the film or EPID image can give quantitative information for calibration of the MLC. If this test is done in an absolute way, the distance between leaves of a leaf pair is verified and the absolute position of the leaf can be determined as shown in Figure 4.1. The measurement accuracy of this technique can be of the order of 0.2 mm (1SD).
Instead of using film, linear (1D) or 2D arrays such as an EPID (Baker et al., 2005) are used for the verification of leaf position. Figure 4.2 shows as an example the results of measurements made with a linear array of 47 ionisation chambers having an area of 4 mm x 4 mm and a distance of 8 mm. By performing the measurements at 80 cm SSD each leaf matches one detector. The measured data were used to adjust the leaf positions more accurately for IMRT applications of the linac after it was accepted earlier with larger leaf position tolerances.

**Figure 4.1** A strip-test design for MLC calibration purposes showing nine adjacent segments 2 cm wide with 1 mm gap, and two extra segments with 4 squares at the left and right side to determine the isocentre, measured with film. Dose profiles are taken for each leaf-pair. The right figure shows the profile of a central leaf. The dose variations of the abutments are used to determine the relative leaf positions, and the measured position of the abutments to determine the absolute leaf position (from Sastre-Padro et al., 2004).

**Figure 4.2** Leaf positions after initial acceptance of the linac and after adjustment for IMRT applications, measured with a linear array of 47 ionisation chambers. (Courtesy Thijs Perik, The Netherlands Cancer Institute, Amsterdam, The Netherlands)
When calibrating the leaf position different choices can be made for defining the position of the leaf relative to its dosimetric properties. This is especially relevant in the case of a rounded leaf edge. Possible choices are:

- The leaf position is the actual position. The advantage of this approach is the clear definition; the possibility of mechanical/optical calibration and the property that opposing leaves can be placed at the same position (closed leaves). The disadvantage is that the dose profile is not directly related to the mechanical properties of the leaf in the sense that the position of the leaf corresponds to a certain dose percentage in the profile.
- The leaf position corresponds to the 50% dose of the profile. This approach is closest to the conventional definition of the field edge. For rounded leaf edges this position is at some distance inside the leaf, which has as a consequence that opposing leaves cannot be at the same position, and a minimum leaf distance has to be defined. The calibration can be done effectively using a water phantom. In abutting fields or segments the dose at the abutting line is not 100%, since there is transmission through the leaves as shown in Figure 4.3.
- The leaf position is based on the best profile in abutting fields. This calibration position is close to approach 2 but a fraction of a mm more inside the leaf. The advantage is that it gives the best dose distribution with abutting segments. This calibration can be done with the strip-test technique described before.

![Figure 4.3 Dose profiles of leaves with rounded leaf ends with different gaps between opposing leaf positions. The calibration of the leaf position is at the 50% dose point. Dimensions are in cm.](image)

It should be noted that the same definition of the leaf position should be taken into account by the treatment planning system applied for IMRT and other applications. As can be seen in Figure 4.3, different definitions of leaf position may lead to large errors in dose distributions of adjacent fields. A similar observation has been made by Cadman et al. (2002), which was traced using the Radiological Physics Center head-and-neck phantom. Knowledge of the leaf position definition in the treatment planning system is therefore of utmost importance.
The effective gap between leaves of a closed leaf pair is quantified by a MLC parameter called the Dosimetric Leaf Separation (DLS) for a Varian accelerator. The DLS has to be introduced into the TPS and quantifies the mechanical leaf gap as well as the photon transmission, which is related to the leaf edge geometry. The garden fence test, in the absolute measuring mode, can also be performed under different gantry angles in order to study the influence of gravitation on the position of the leaves or on the DLS.

For step-and-shoot treatments with many segments the penumbra is an important parameter of the collimation system. Whereas in sliding window treatments effects of an incorrect penumbra calculation blur out over the entire field, in the step-and-shoot mode there are many static segment edges where penumbra errors can cause problems. Dependent on the design of the MLC and the segment sequencer in the planning system, different components cause different penumbra widths: X- and Y-collimators whether or not in combination with leaves, the leaf edges themselves, and both sides of the leaves. In Figure 4.4 an example is given of the penumbra width resulting from these various possibilities.

![Figure 4.4 Penumbra values (80%-20% dose distance) for an Elekta MLC at the indicated positions measured with film. The arrows indicate the positions where the penumbra values were measured. The vertical and horizontal fat lines show the position of the back-up (Y) and X-collimators, respectively.](image)

<table>
<thead>
<tr>
<th>1: Leaves only:</th>
<th>5.8 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>2: Back-up (Y) collimator only:</td>
<td>4.8 mm</td>
</tr>
<tr>
<td>3: Leaves + back-up (Y) collimator:</td>
<td>4.2 mm</td>
</tr>
<tr>
<td>4: Side of leaves (left):</td>
<td>3.8 mm</td>
</tr>
<tr>
<td>5: Side of leaves (right):</td>
<td>3.5 mm</td>
</tr>
<tr>
<td>6: X-collimator only:</td>
<td>3.6 mm</td>
</tr>
<tr>
<td>7: X-collimator + leaves:</td>
<td>3.4 mm</td>
</tr>
</tbody>
</table>
Figure 4.5 Film measurement of an IMRT field delivered using the sliding window technique of a head-and-neck treatment plan transferred to a phantom. The measured and calculated dose distributions along the red line have been compared. The correct value of the DLS parameter for this set-up was 2.6 mm. With this value the calculated and measured data agreed very well and are all within gamma criteria of 3% local dose difference and 2 mm DTA. The calculations were repeated by using a larger DLS of 3.1 mm. As a result 9% of the area inside the 0.14 Gy isodose area had a gamma value larger than 1.

Other leaf parameters determining the dose delivery of an IMRT treatment are, for instance, the inter-leaf transmission and the tongue-and-groove effect. To facilitate the use of MLCs as moving mechanical devices, they have a stepped-edge or tongue-and-groove design. Consequently when the radiation passes first through the tongue part in one segment and then through the groove part of the leaf in an abutting segment, in the junction area there will be an under-dosage of adjacent leaves, which may amount up to 25% (e.g., Essers et al., 2001, Webb, 2006). It should be noted that not all available TPS consider the tongue-and-groove effect. While these characteristics are constant for a specific type of MLC, and most likely do not change with time, the DLS should be tested frequently because it can change for instance by a wrong calibration of the leaf position. In Figure 4.5 dosimetric differences are shown in case the DLS differs by 0.5 mm.

4.1.2 Step-and-shoot treatments

IMRT treatments based on the step-and-shoot principle include usually several segments with a limited number of monitor units, e.g. less than 10 MUs. When segments like these are used, it is essential that the performance of the accelerator is within the required tolerance as soon as possible after start-up. During start-up, depending on the accelerator design, the energy can differ shortly from the intended energy. This can have an effect on all kinds of beam characteristics like dose per MU, PDD and beam profiles. It is important to determine the start-up characteristics during commissioning and to check these regularly in the QA
programme. In Figure 4.6 an example is given of the influence of the magnetron design on the beam calibration for a limited number of MUs. It is generally assumed that modern accelerators have a stable output after delivering a few MUs.

Figure 4.6 Beam calibration for a limited number of monitor units depending on the type of magnetron and steering technique for Elekta accelerators. In 1997 the feedback technique with slits was used. An improvement of this technique was the slitless flight tube, which was followed by a new design magnetron with faster tuning (Courtesy Geoff Budgell, Christie Hospital, Manchester, UK).

### 4.1.3 Sliding window techniques

Quality control programmes of linacs applying sliding window techniques for IMRT delivery are generally based on the tests developed by LoSasso et al. (1998) and Chui et al. (1996). A summary of the actual application of these tests, or with some modifications, in 5 institutions applying the sliding window technique, has been given by Van Esch et al. (2002). Measuring leaf positions for the sliding window technique can be done by creating a uniform field of 10 cm x 10 cm with a sweeping gap of 5 mm (LoSasso et al., 1998). By measuring the signal of an ionisation chamber placed on the beam axis, a positional difference of 0.1 mm can be detected in this way. These measurements can be performed with the same phantom-ionisation chamber combination as applied during the routine output measurements of the accelerator but should be performed under different gantry angles to assess the effect of gravity on leaf positioning.

Several groups have reported verification of leaf position during sliding window dose delivery using EPIDs. A method to derive leaf positions in each camera frame and to compare them with a table of prescribed positions has been presented by the Royal Marsden group using a video camera based type of EPID. The spatial and dosimetric resolution of the sy-
stem permits verification to within the limits of ± 3% dose or ± 3 mm displacement of an isodose line, whichever is lower (Partridge et al., 2000). The system has been used for the pre-treatment verification of IMRT of breast treatment. Ma et al. (1997) calculated reference images from MLC leaf sequencing files and compared these with images measured with a fluoroscopic beam imaging system. Under-dosage due to the tongue-and-groove effect, up to 30%, has also been measured with EPIDs (Essers et al., 2001). Vieira et al. (2002) developed a method based on a flat field produced with a 5mm-wide sliding gap for each leaf pair. Deviations in gap widths are detected as deviations in grey scale value profiles derived from the EPID image, and not by expressing leaf positions in the images. It was shown that errors in leaf gap as small as 0.1 – 0.2 mm could be detected in this way. Recently this method was adapted for step-and-shoot treatments (Vieira et al., 2006). In the same paper these authors discussed causes for deviations between actual and prescribed leaf positions.

For the sliding window mode it is also necessary to check whether the leaf system can still perform the maximum estimated speed that the leaf motion calculator takes as basic parameter in order to calculate the leaf motion. In the “speed test” reported by Chui et al. (1996) all leaf pairs move with a constant but different speed, generating a stepwise dose delivery of well-defined intensity levels. The effect of beam interrupts on dose delivery can be assessed using the same test. The influence of beam interruption should also be checked for a complete IMRT field by measuring the dose distribution when periodic beam stops during delivery are introduced. These measurements have to be compared with uninterrupted data and no dosimetric differences should be measured.

4.2 OTHER TYPES OF IMRT DELIVERY SYSTEMS

Historically, IMRT delivered by tomo(slice)therapy machines preceded that by MLC-based techniques, as discussed by Webb (2006) in his review of IMRT delivery techniques. The first commercially available serial tomotherapy machine, designed by the NOMOS Corporation (North American Scientific, Chatsworth, CA, USA), applied a static approach in which the patient and couch did not move during gantry rotation while the beam was on. A single slice is treated per revolution of the gantry followed by a longitudinal movement of the table to the next slice. This is repeated until the whole length of the target volume is covered. The radiation is collimated by means of two sets of 20 leaves that can either be in or out during gantry rotation. The whole collimator system can be fit on an existing accelerator and is supported by its own treatment planning system (PEACOCKPLAN). Helical tomotherapy is similar but the patient is slowly translated through the radiation field while the gantry is rotating. Serial tomotherapy has not been applied at a large scale in Europe and QA of such a device will therefore not be discussed in this booklet. Helical tomotherapy machines have been, or are in the process of being installed at a number places in Europe (Beavis, 2004) and the specific QA aspects of those machines will therefore briefly be discussed.
Several other techniques have been developed to deliver IMRT as discussed by Webb (2006). The most flexible system is the CyberKnife, designed for stereotactic radiotherapy, delivering pencil beams of photon beams of different fluence generated by a linear accelerator mounted on a robotic arm. QA procedures relevant for its use as an IMRT system will also briefly be discussed in this chapter.

4.2.1 Helical tomotherapy machines

Helical tomotherapy has been developed at the University of Wisconsin over several years (Mackie et al., 1993, 1999), and clinical machines are now manufactured by Tomotherapy Inc. (Madison, WI, USA). The equipment is designed to be a purpose-built image-guided radiotherapy system, to deliver IMRT whilst also being able to verify the set-up of the patient and, in the future, the dose delivered during treatment. The helical tomotherapy machine can be described as a combination of a helical CT scanner with a linear accelerator. A compact linear accelerator is mounted on a rotating gantry at a source-axis distance of 85 cm, generating a nominal 6 MV photon beam. Differently from a conventional linac, the device does not include a physical flattening filter, and consequently, its dose rate is increased to about 800 cGy/min. The primary fluence emanating from the treatment head is forward peaked until it is modulated by the MLC. The longitudinal (y-direction) field width (slice thickness) is defined by a pair of moveable jaws, able to define an open maximum field size equal to 40 cm (x) x 5 cm (y). A slice thickness of 1.0, 2.5 and 5.0 cm can be chosen. Laterally (x-direction), the beam is collimated by a binary MLC, consisting of 64 interdigitised leaves with a nominal width of 6.25 mm at the isocentre. The leaves operate in a binary way: completely opened or closed in 20 ms, turning individual beamlets on or off. Intensity modulation is accomplished by varying the fraction of time for which different leaves are opened. The gantry rotates at a constant velocity with a period between 10 s and 60 s; clinically a period longer than 15 s is always set for patient treatments.

The same radiation source, detuned to a nominal 3.5 MV photon beam, is also used for megavoltage image acquisition: a xenon-detector array mounted opposed to the radiation source allows 3D tomographic reconstruction of the body structures in order to check and to correct the patient’s set-up. Volumetric image acquisition can be performed at dose values comparable with those required to obtain images with conventional MV electronic portal imaging devices; i.e., of the order of 2 cGy.

Helical tomotherapy is capable of delivering conformal dose distributions using 51 beam projection angles and some 30,000 equivalent segments. Verification of delivering these highly complex treatments is therefore extremely important and QA procedures have been developed to ensure the safe and consistent operation of a helical tomotherapy facility. Accurate dose delivery depends on various factors related to: 1) mechanical and geometrical
characterisation of a Tomotherapy unit; 2) the dosimetric beam configuration, and 3) the system dynamics and system synchrony. These require special QA tests, which will briefly be discussed as far as they are different from those of conventional linear accelerators (e.g., see Balog et al., 2003b, Broggi et al., 2008). The latter authors recently reported the results of a two-year QA programme for the physical and dosimetric characterisation of their Tomotherapy unit.

1) The mechanical and geometrical characterisation of a Tomotherapy unit, as for a conventional linac, deals basically with the alignment of all system’s components: lasers, source, moveable jaws, multileaf collimator (MLC), detector array. Different tests are included as discussed in detail by Balog et al. (2003a) and concern the verification of, for instance, MVCT alignment with respect to radiation source and moveable jaws, and couch drive distance levelling and alignment. The impact of these parameters on the dose distribution delivered to a patient is similar to that of a conventional linac with some peculiar differences, mainly related to the accuracy of laser alignment. Because the patient cannot be directly positioned inside the bore, the lasers point to a “Virtual Isocentre” set outside the bore at a fixed distance (70 cm) horizontally away from the actual radiation isocentre. Patients are set up at this Virtual Isocentre and then the couch shifts to the radiation isocentre. For a correct and precise patient positioning and treatment, fundamental prerequisites are the accuracy of the Virtual Isocentre position, with respect to the radiation isocentre inside the bore. Secondly, accurate couch drive levelling and alignment is even more critical for helical tomotherapy than for conventional therapy. Test patterns to check MLC position, alignment, leakage, and timing of a Tomotherapy unit have recently been described by Sarkar et al. (2007).

2) The dosimetric beam configuration requires measurements similar to those of a conventional linac concerning beam output, energy, lateral and longitudinal profiles, output reproducibility and linearity, but with some small differences related to:

- **Machine output.** Tomotherapy does not work on a monitor unit-based system and the output is therefore calibrated in terms of reference dose-rate (cGy/min) at a specific depth in water for a fixed field-size with the SSD set at SAD.
- **Output factors.** The field-size dependence of the output can be adequately modelled based on three components. First, the head scatter variation for different jaw settings; secondly, the phantom scatter variation, which is taken into account by the convolution-based dose calculation algorithm of the planning system, and thirdly, the tongue-and-groove effect, which is also calculated by the treatment planning system.
- **The off-axis profile in the lateral direction (x- or cone-profile) and in the longitudinal (y-) direction.** Tomotherapy machines do not have a flattening filter, and therefore the penumbra width is mainly determined by the forward peaking of the primary radiation beam. This is reflected in the shape of the lateral cone-profile, while the couch
movement determines the shape of the y-profile. Classical measurements of the off-axis dose distribution, such as field flatness and symmetry, are not very relevant. On the other hand, the helical nature of the irradiation delivery makes periodical acquisition and check of the longitudinal profiles crucial in order to avoid an unplanned treatment over- or under-dose situation. It is therefore important that the profile shapes measured at commissioning and modelled by the Tomotherapy treatment planning system accurately reflect machine performances at the time of treatment (see Figure 4.7). The use of a commercial diode array for QA of these characteristics has been proposed (Langen et al., 2005). In Sections 6.1.5 and 6.1.8 more details will be provided of patient-specific QA procedures applied for helical tomotherapy treatments.

- **Depth dose curves.** Due to the reduced source-axis distance of 85 cm, depth dose curves of a Tomotherapy unit have a steeper dose fall-off than those of a conventional linear accelerator (see Figure 4.7).

![Figure 4.7 Transversal cone-shape profiles and PDD curves of a Tomotherapy unit.](image)

The main peculiarity of a Tomotherapy unit is to provide IMRT helical irradiation patterns, thanks to the concomitant and synchronised continuous gantry rotation, couch translation and opening/closing MLC leaf movement. For this reason, a precise characterisation of the synchronised components is necessary together with an accurate couch speed and couch movement check.

3) QA of the **system dynamics and system synchrony** aspects of a Tomotherapy unit concerns tests similar to those applied for verifying sliding window IMRT techniques although with some specific peculiarity (Fenwick et al., 2004). The total dose delivered at a point of a target volume varies linearly with the length of time for which the point lies within the radiation beam. Tests in this category concern the verification of treatment beam characteristics such as:

- **Width of the field** in the y-direction (defined by the jaws).
- **Couch velocity.** The time that a point within the patient’s target volume lies in the radiation field varies linearly with the width of the field and inversely with the couch speed.
• **Leaf latency effects.** The dose delivered to the patient scales linearly with the actual fractional open time, which differs from the programmed open time due to “leaf latency” effects, the finite opening and closing time of the leaves. This phenomenon is taken into account in the dose calculations performed by the treatment planning system (Balog *et al.*, 2003b). Latency corrections are determined during commissioning by measuring the energy fluence when a leaf is opened as a function of fractional programmed opening time and projection time.

• **Synchronisation effects.** Specific synchronisation components have also to be tested during the characterisation of a Tomotherapy unit to match a delivered dose distribution with the planned one. These tests concern the synchronisation of gantry angle with leaf opening, with couch drive and with linac pulsing.

Due to the complexity of the treatment approach, it is desirable to have a simple tool for pre-treatment verification of each helical tomotherapy treatment. For this purpose a special phantom has been designed for tomotherapy treatment verification, the so-called “cheese phantom”, (see Figure 4.8). The cheese phantom is a cylindrical solid-water phantom of 15 cm radius and 18 cm length cut into 2 semi-cylindrical halves with a piece of film sandwiched between the two halves. Simultaneously single or two ionisation chamber measurements are performed at points along an axis perpendicular to the film, using an ionisation chamber having a 0.056 cm³ volume and a wall thickness of 1.1 mm; both wall and central electrode are made of C552 air-equivalent plastic. Some results of its application will be given in Sections 6.1.5 and 6.1.8. More recently Van Esch and co-workers proposed an on-line method using a 2D ionization chamber array (Seven29, PTW, Freiburg, Germany), discussed in Section 3.3, inserted in a dedicated octagonal phantom (Van Esch *et al.*, 2007). To compensate for the directional dependence of the 2D array, built-in cylindrically symmetric compensation cavities were added.

![Figure 4.8](image_url)  
*Figure 4.8* Solid water “cheese phantom” used for patient-specific QA of helical tomotherapy treatments showing the position of the ionisation chamber, inserts and film.
Radiographic film placed under the patient, exposed from most of the possible beam directions, might constitute a useful in vivo dosimetry record of a helical tomotherapy treatment. Measurements were performed during the initial clinical implementation of helical tomotherapy in London, Ontario, Canada, on all patients during the first fraction (Kron et al., 2005). The optical density of the film could be determined using a dose calculation on a phantom of similar size as the patient. The comparison of expected and delivered dose allowed the verification of dose delivery patterns, which was found to be particularly useful in the case of treatment interruptions. The dose measured with film differed in general less than 10% from the expected value despite the fact that no build-up material was used on the film. The agreement improved with proximity of the target volume to the location of the film on the treatment couch. Due to the rotational delivery mode, radiographic film was shown to be a useful, cheap and convenient method to verify dose delivery in helical tomotherapy.

Another approach for clinical delivery verification and dose reconstruction during helical tomotherapy has been described by Kapatoes and co-workers (2001). The energy fluence is used in conjunction with an image of the patient in the treatment position, obtained with the CT detector system, to reconstruct the full three-dimensional dose distribution. For this purpose an extensive database of detector signals is used, obtained by means of a large number of phantom measurements. The detector signals are collected as a function of two parameters: radiological path-length and detector-to-patient thickness. Both are obtained from a CT image taken at the time of delivery. Their method was applied for the verification of a simulated prostate cancer treatment and a nasopharyngeal delivery on a dog cadaver. For both cases, it was found that the verified energy fluence and dose results using the data base approach agreed very well with those using other methods. Interestingly in their publication the authors noticed a severe under-dosage of the target volume using their method. This error was caused by a rotational offset present in the delivery due to a lack of synchronisation between the rotating and translation system and the rest of the delivery system.

### 4.2.2 CyberKnife systems

The CyberKnife stereotactic radiosurgery system was developed by Accuray Inc. (Sunnyvale, CA, USA) and is a device designed for robotic image-guided stereotactic radiosurgery and radiotherapy throughout the whole body. Initially the CyberKnife was developed to overcome the shortcomings of frame-based stereotactic treatment with the Leksell Gammaknife and dedicated linacs. The Gammaknife has inherent mechanical limitations that exclude treatments at extreme locations in the skull and extracranial lesions. Dedicated linacs have the inherent mechanical limitations of the gantry. The improved maneuverability of the CyberKnife linac by the robotic arm and the use of circular pencil beams realise the dose coverage of the tumour volume by “dose painting” (Ling et al., 2000), which can be considered as a “spatial” IMRT technique. The main CyberKnife components are:
• A 6-axis industrial robotic arm as a high-precision positioning device providing a spatial accuracy of ± 0.2 mm in every reachable position.

• A compact 6 MV linear accelerator delivering 400-600 cGy/min at 80 cm SAD. During the treatment procedure the robot moves the linac from one radiation position (node) to the next. A treatment plan consists of 50 up to 200 single beams or more and the entire treatment time ranges from 20 minutes up to one hour or more in the case of complicated clinical situations.

• An integrated stereotactic X-ray system for target tracking consisting of two orthogonal X-ray tubes, rigidly mounted to the ceiling, and two corresponding amorphous silicon imagers installed on the floor of the treatment room.

• An automated treatment couch offering 5 degrees of freedom for the necessary translational and rotational couch corrections.

Tumour tracking is based on tumour position relative to the bony structure (e.g., the skull) or implanted gold fiducials. Directly prior to treatment, the CyberKnife system establishes a correlation model between LED markers attached to the patient’s chest and the implanted fiducials. During the treatment the outer markers are continuously tracked by a camera system and the robot motion is synchronized to the target movement by using the correlation model, which relates the outer and inner movement. Similar systems applied for tumour tracking using conventional linear accelerators have been described elsewhere (e.g., Mageras et al., 2006). QA of 4DIMRT techniques is a relatively new field as mentioned in Section 2.1 and will not be discussed in detail in this report.

The CyberKnife system offers a combination of a non-isocentric and non-coplanar beam delivery with a nearly real-time image guidance allowing tracking of the target or surrogate continuously during radiation delivery. The impact of these two features make the CyberKnife treatment procedure quite different from other IMRT procedures, both in clinical and in technical aspects. As there is neither an isocentre nor a single treatment plane applied, the treatment planning system works by optimising about 1200 beams in order to get the optimum final treatment plan. A summary of the characteristics of the CyberKnife system, a description of the adaptation of the system to treat extra-cranial tumours and the way the equipment is clinically applied can be found elsewhere (e.g., Adler et al., 1999, Romanelli et al., 2003, Webb, 2006).

Due to the highly flexible beam delivery geometry and the lack of MLCs, the QA programme of the CyberKnife differs from that of conventional linacs. The following parts are of a QA programme, specific for the CyberKnife (CRCPD, 2006).

• **Daily:** system status, safety interlocks, output calibration

• **Monthly:** beam parameters, robot mastering (visual), imaging targeting and imaging alignment, film - phantom target test
• **Quarterly:** target locating system tracking, linac laser, mechanical and radiation field alignment (the CyberKnife uses a point laser that is coincident with the radiation field as a QA tool. The laser is reflected by an adjustable mirror and its alignment must be checked periodically)

• **Annually:** beam commissioning, robot mastering (electronic), and couch indexing accuracy.

Typical CyberKnife applications are the treatment of intracranial lesions close to critical structures that require small margins (≤ 2mm). To establish the accuracy of this robotic treatment a variety of methods for in-phantom dosimetric verification of dose distributions with clinically realistic and well-quantified movements during irradiation are designed by the users. The aim is to verify the dose distribution in the TPS after commissioning, especially in areas with steep dose gradients, as a check of the beam data (output factors, beam profiles and tissue-phantom ratios of small beams). One of the QA tools delivered with the CyberKnife is a special human-like anthropomorphic head phantom with a film-cube insert for film measurements with radiochromic EBT film. Figure 4.9 shows a simple CyberKnife treatment plan using the 20 mm diameter collimator to deliver a dose of 4 Gy at the 100% isodose surface encompassing a 30 mm target. A critical structure was mimicked to create a steep dose gradient (80%-20% dose reduction in 5 mm). The figure also shows that good agreement exists between the measured and calculated dose distributions after irradiation with well-defined phantom translations and/or rotations during irradiation.

**Figure 4.9** Left: A “skull tracking” TPS plan (i.e., tumour tracking based on the bony structure of the skull). This plan uses 69 beams with a 20 mm diameter collimator. Right: 2D dose patterns of skull tracking experiments (film size: 6.35 cm x 6.35 cm). Intentionally inflicted translations and rotations during phantom irradiation were satisfactory corrected by the robotic system.
4.3 TREATMENT PLANNING SYSTEMS

Several reports exist giving recommendations for commissioning and QA of treatment planning systems as discussed in Section 1.1. In these reports test cases are described to be used for systematic verification of the accuracy of dose calculations performed for specific well-described situations. Numerous geometries might in principle be interesting to be investigated, as for instance elucidated in IAEA Report TRS 430 (IAEA, 2004). Attempts have therefore been made to reduce the number of tests to the most crucial ones, as proposed for instance in ESTRO Booklet No.7 (Mijnheer et al., 2004) and more recently by Camargo et al. (2007). However, all these tests are basically designed for treatment planning of 3DCRT. Most of these tests are therefore a good starting point for commissioning a TPS that is also going to be used for IMRT. Some of these geometries, for instance, complex MLC-shaped fields, are useful test cases but not sufficient when implementing IMRT in the clinic. Therefore testing dose calculations of fields designed by a TPS should be extended to IMRT fields of various shapes including small fields. In principle this can be done by comparing the dose calculations performed by the clinically applied TPS with experimentally determined dose values, e.g., by using small volume detectors or films. If needed, commissioning of a TPS needs to be extended towards small fields. Because the planning and delivery of IMRT fields are intimately connected, verification of a TPS should preferably be performed for the treatment machine for which it clinically will be used.

At this moment only very few publications (e.g., Van Esch et al., 2002 and Sharpe, 2003) or reports (e.g., Ezzell et al., 2003) exist describing test situations to be investigated before starting with IMRT in the clinic. The current approach is that centres that start clinically with IMRT design their own test programme, based on the possibilities of their TPS, the intended IMRT indications, the complexity of the related clinical IMRT techniques, the advice and support of other users of the same system, as well as the confidence in their own approach. Some of these tests have, however, to be performed by all users of a specific TPS to be sure that no errors have been made, for instance, in the beam modelling of the TPS, and to guarantee in general the safe introduction of new IMRT treatment techniques. These tests may include verification of:

- definition of leaf positions in the TPS;
- beam profiles of small segments, as well as of abutting fields, during step-and-shoot dose delivery;
- beam profiles of small fields created by sliding window techniques, including verification of the interplay of leaf speed with dose rate;
- the tongue-and-groove effect and the leaf transmission;
- output factors and depth dose curves of small square fields, i.e. smaller than the fields verified during the acceptance of the system for 3DCRT;
- output factors and depth dose curves of asymmetric fields of different shape, including...
small elongated fields, having different distance and over-travel with respect to the central axis;

- dose distributions in inhomogeneous phantoms irradiated with small fields;
- dose distributions of some typical site-specific fields having segments characteristic for that treatment technique;
- a number of test patients for each treatment site, e.g., prostate, head-and-neck, lung, breast.

Measurements in small fields require special precautions. Dosimetry systems commonly used for IMRT verification have been discussed in Chapter 3. A number of the tests mentioned above can be performed with the dosimeters discussed in that chapter: small ionisation chambers, film, two-dimensional arrays or EPIDs. Although not suitable for absolute dose determination, other small volume detectors (diodes, diamond detectors) are useful for relative measurements in small fields (e.g., Aspradakis et al., 2005 and McKarracher and Thwaites 2006).

It should be noted that at this moment no acceptance criteria are available for the tests of a TPS specific for IMRT applications. This situation is different from that of treatments using 3DCRT techniques. Reports dealing with commissioning and QA of a TPS until now had clearly defined pass-fail criteria for well-described tests using a large variety of geometries, thus providing clinical physicists guidelines to accept specific aspects of a TPS. For IMRT the situation is different, which is due to a number of reasons. First of all the combination of a large variety of segments having different shape creates IMRT fields of completely different intensity profiles, which are difficult to compare. Consequently it is virtually impossible to define recommendations for the accuracy of dose calculations of constituting individual segments. Furthermore, the accuracy requirements might be dependent on the dose level applied in IMRT. Finally, the interaction of dose calculations performed by a TPS and the performance of the treatment machine makes it very difficult to separate the two components contributing to the overall uncertainty in the dose delivery of an IMRT treatment. For example, the mechanical constraints of a certain type of MLC have direct impact on IMRT delivery, i.e. the sequencing of fluence patterns. Thus a specific TPS might require different QA tests depending on with which treatment unit it is combined. For those and other reasons the emphasis on QA of IMRT is shifted from acceptance testing and commissioning of a TPS to patient-specific QA. However, also for patient-specific QA no guidelines are available. In Chapter 6 the current clinical experience in Europe is described giving an indication how this situation is solved in various institutions having ample clinical experience in IMRT.
5. INDEPENDENT DOSE CALCULATIONS APPLIED FOR IMRT VERIFICATION

There are three levels of verifying the accuracy of dose calculations performed by an IMRT treatment planning system: (a) by the manufacturer of commercial systems, (b) in the clinic during the acceptance testing and commissioning of the TPS and the pre-clinical phase of IMRT, and (c) as part of a patient-specific quality assurance programme. At present, experimental verification has become the preferred method for IMRT commissioning as well as for patient-specific QA. During such procedures typically phantom measurements of the composite treatment plan or single field verification are carried out. Although there are ways to optimise the workflow of such experimental methods, as elucidated in the next chapter, experimental treatment plan verification has become the bottleneck for clinical IMRT implementation in many departments. In this chapter we will discuss other methods that may lighten the burden of experimental verification by introducing calculational methods instead. One method, which shows great promise, is to verify the complete 3D dose calculation performed by the treatment planning system using a Monte Carlo dose calculation engine. Such an approach is still rather time consuming. An independent dose calculation at one or more points, which is often recognised as an appropriate QA tool for conformal radiotherapy, may also be adequate for IMRT verification in combination with a comprehensive QA programme of linac performance. Both approaches will be elucidated in this chapter.

5.1 MONTE CARLO CALCULATIONS

Dose calculations performed by commercial treatment planning systems have their limitations, for instance at interfaces of tissues having different density and composition, or for small and irregular fields. These problems may be magnified when using IMRT techniques due to the high complexity of these treatments in which often a considerable amount of small beam segments is involved. Furthermore, the head scatter, the leaf ends, the leakage through the leaves and the tongue-and-groove effect are difficult to model and to take into account in an accurate way. Monte Carlo (MC) dose computation methods allow in principle the highest accuracy since its results are only limited by the accuracy of the underlying interaction cross sections and the way they are used. From this argument stems the frequently held view that the greatest advantage of MC is its precise consideration of particle transport in geometries of highly variable shape, density and composition. However, for the application of MC in commercial treatment planning systems, this view is too simplistic as it only emphasizes the passive particle transport in the patient, yet neglects the generation of particles and photons in the linear accelerator. Hence, the accuracy of MC calculations for IMRT verification is essentially a question of the model of the particle sources in the accelerator head (Reynaert et al., 2005). On the other hand, the elegance of the MC method when dealing with complex
geometries comes to bear most fruitfully in the consideration of the various effects created by the beam defining collimators (MLC), such as photon scatter, variable transmission, leakage, and curved focussing edges of the leaves. Here it is important to notice that the model of the invariable, beam producing components of the accelerator (electron beam, bending magnet, target, primary collimators, filters) can be fully separated from the model of the variable components (collimator jaws, collar adjusting leaves, wedges, compensators), and the patient model. This is usually not possible for accurate pencil beam dose computation methods, where energy kernels and fluence in the patient need to be corrected for field size and tissue heterogeneity. As a consequence, the usefulness of MC calculations for IMRT verification rests on the fact that for a proper particle source model, the accuracy of the dose computation does not depend on the field shape (for the same primary fluence, i.e. number of particles per area impinging onto the patient) while also tissue heterogeneity is directly taken into account. The usefulness of MC calculations as an independent check of dose calculations has been demonstrated in several studies including those of Ma et al. (2000) and Leal et al. (2003). In the following paragraphs, various approaches to build particle source models and variable component models are presented. Particular attention is paid to commissioning the model for a given clinical accelerator. It should be noted that the final accuracy in the calculated dose distribution will never be better than the accuracy of the beam source model, i.e. the modelling of the treatment head with all its components, which is one of the fundamental limitations of MC calculations for IMRT verification.

For the simulation of accelerator head models, widely used utility packages, such as BEAM, which build on established MC codes, e.g. EGSnrc, can be employed (Rogers et al., 1995, Briesmeister 1997, Rogers et al., 2003). A problem that sometimes arises is that manufacturers may not supply detailed information about all materials used in their head construction, for instance the % tungsten in MLC leaves, thus introducing additional legitimate sources of uncertainty in MC codes. While the components of the accelerator can in principle be described in great detail, the properties of the electron beam on collision with the target are harder to capture. The relationship between these properties like energy spectrum, spot diameter, angular and spatial distribution and measurable dosimetric quantities of the photon beam, such as beam profiles and depth dose curves, is well understood (Sheikh-Bagheri and Rogers, 2002). However, tuning these parameters for a given clinical accelerator is still a cumbersome process, which requires many iterations of parameter adjustments and simulation runs. Experimental verification of a specific beam delivered by a linac simulated by MC is then generally performed by comparing depth dose curves for various field sizes and beam profiles at various depths as shown in Figure 5.1. Once the parameters have been set, a phase space file that contains momentum and location of a large number of particles in a plane above all variable components can be recorded and used for subsequent dose computations.

Alternatively, the essential properties of an accelerator can be captured in a virtual source model, consisting of various photon and electron sources characterised by a certain energy
spectrum, diameter and angular distribution. The great advantage of such a type of model is that its parameters can be determined directly from measurements in air and water (Fippel et al., 2003). Given the uncertainties associated with the full simulation starting with the electron beam, these models are in many situations in no way inferior, as for instance is shown in Figure 5.2. Good agreement is found between measurements and a virtual source model in Hyperion (XVMC). The mean differences are on average 0.5% and locally 2% at maximum for the depth dose curve and inside the beam profile, and out-of-field locally 3% at maximum.

The collimators can be modelled explicitly, which requires both highly detailed MC simulation modules and significant computation times (Van de Walle et al., 2003). This method takes also secondary particle generation in the collimators into account. Alternatively, the collimators could also be modelled as transmission filters, which can include some detail like inter-leaf leakage, but cannot yet accurately model scattered particles (Wang et al., 2002). However, this is a small price to pay for a significant speed increase, because the MLC scatter contribution becomes smaller for decreasing field sizes (Kim et al., 2001), and can therefore be described as a diffuse particle source in the accelerator head, while its overall contribution to the dose is of the order of a percent.
Figure 5.1 Measured and calculated depth dose curves and beam profiles for various field sizes to assess the nominal energy for the simulated MC source corresponding to a 6 MV beam.
To tune the particle source model to a given accelerator, it is usually sufficient to measure reasonably sized beams in air and water. The primary particle spectrum can be obtained from depth dose curves of small fields of for instance 5 cm x 5 cm, which are small enough to block out most head scatter, yet big enough to avoid systematic measurement errors. Unless the dose model is to be used for stereotactic fields of less than 2 cm x 2 cm, it is not necessary to measure these fields for commissioning. The spot size of the primary photon source begins to influence the output factor only for even smaller field sizes. In this case, it is sufficient to measure the output for a number of small fields to fine-tune the source diameter (Sikora et al., 2007). This greatly reduces the complexity of the commissioning measurements and the impact of systematic errors related to dose measurements in small fields. In order to test the model for IMRT, a test case like in Figure 5.4, which features a medium-sized IMRT field at large off-axis distance, is particularly relevant. It turns out to be one of the most decisive tests for the correctness of such factors as photon energy spectrum, source diameter and off-axis spectrum softening to verify output factors for small fields with measurements.

The variation of output factor with field size is a consequence of the imperfect collimation of a blurred photon source in the accelerator head. Closely related is the dose outside the field, which stems only partially from leakage but primarily from photons scattered in the flattening filter and primary collimators. Note that by the nature of IMRT delivery with an MLC, the ratio between direct and indirect irradiation of the target volume, and especially of organs at risk, is much smaller than for other types of conformal RT (e.g., see Schwarz et al., 2003).

Despite great advancements in computer hardware and MC code development, computation time can still become limiting. In particular, reducing the voxel size of the dose simulation grid by half may increase the computation time, at the same level of statistical accuracy, by a factor 64! Hence, running a verification computation for each sub-field of an IMRT delivery on a 1 mm grid to a low statistical uncertainty is still a formidable task. Affordable multi-core
CPU technology will boost performance considerably in the near future as MC can take full advantage of parallel computing. Although coarser simulation grids may blur the fine detail of the dose distribution, they may be closer to the measurements, where most detectors, apart from film and some types of EPIDs, have relatively large effective volumes. Ideally, the detector should be included in the MC simulations of the phantom, and the local dose at the site of the detector corrected for the detector response, as shown in Figure 5.3 (Dohm et al., 2005). This is especially appealing for film measurements, where the spectral dependence of the film can now be included in the simulations as the local energy spectrum can be determined by simulations (Palm et al., 2004).

![Figure 5.3](image-url) Field-size dependence of output factors for fields centred 10 cm off-axis. Both BEAMnrc and XVMC simulations are corrected for the size of the ionisation chamber. The magnitude of this correction becomes visible in the uncorrected XVMC results.

For the purpose of IMRT verification, the elegance of MC calculations stems from the fact that for a proper source model, the accuracy of the dose computation is virtually independent of the field shape. This offers the opportunity to commission the source model with dose measurements with small systematic error, and to compute the dose for fields where measurements are difficult. In the example shown in Figure 5.4, a very patchy fluence pattern was delivered which means that most parts of the field were shielded for the majority of the 25 segments and exposed by only a few. Still, most points along the line are consistent with a MC dose computation within a 2mm/2% criterion. It should be noted that the tongue-and-groove effect is modelled in the MC calculation and the absolute dose is calculated. Film was
not modelled, neither by density nor in terms of response. In Section 6.1.3 more details of this specific example will be given, whereas in Section 6.1.4 an example will be given of the use of MC dose computation for patient-specific QA purposes.

MC calculations are not only used for the verification of the solution supplied by a commercial planning system but can also be applied to obtain the shape of the beam segments and the corresponding number of monitor units during the optimisation process, i.e. the leaf sequence (Jeraj et al., 1999). However, such a procedure cannot use full MC calculations because the task to find the best set of segments is an iterative process. Therefore, to obtain the solution in real time it is necessary to use approximations in the transport and geometry as demonstrated, for instance by Ma et al. (2000) and Leal et al. (2003).

The future of MC tools for IMRT verification may see further speed increases and more detailed collimator simulations. Furthermore, the inclusion of the detector into the simulations in order to compute the detector signal instead of dose may reveal several limitations of currently applied tools for the verification of IMRT, as discussed for ionisation chambers in Section 3.1. Eventually, MC dose computation in treatment planning may even play a role to reduce the effort of plan verification altogether, as elucidated in Chapter 7. Recently a European Workgroup on Monte Carlo Treatment Planning (EWG-MCTP) has been created to offer a platform to scientists to exchange information, to develop new ideas and to initiate international collaborative programmes. The proceedings of The First European Workshop on Monte Carlo Treatment Planning organised by EWG-MCTP in Ghent, Belgium has been published in the online journal JPCS (Journal of Physics: Conference Series, http://www.iop.org/EJ/conf).

Figure 5.4 Left: 6 MV step-and-shoot IMRT field measured with film. Right: dose profile along the green line in the left image resulting from a XVMC dose computation and film (Kodak X-Omat V).
It should be noted that at this moment only few institutions are using Monte Carlo successfully in their QA programmes, but these institutions are typically doing development work in this area. Future advances will be required to make this method of IMRT verification more practical. In addition there needs to be close collaboration with equipment manufacturers to get all necessary information, or the phase space above the patient beam modifiers. In summary, Monte Carlo methods are very promising for numerous applications including QA of IMRT, but will require more development work and improved input, in addition to speed, before they become commonly used.

5.2 OTHER METHODS

In three-dimensional conformal radiotherapy (3DCRT), based on uniform intensity beams, independent dose calculation/monitor unit (MU) verification is often applied as a patient-specific QA procedure. During the last decade three documents have been published giving recommendations for MU determination and/or MU verification: by ESTRO (Dutreix et al., 1997, Mijnheer et al., 2001) and by the Netherlands Commission on Radiation Dosimetry, NCS, (van Gasteren et al., 1998). However, none of these recommendations covers independent MU calculations for IMRT. In 3DCRT persons having ample experience in treatment planning can estimate the number of MUs that a treatment field requires. On the other hand, in IMRT it is hard to intuitively evaluate the number of MUs needed to deliver a certain dose with a dynamic or static intensity modulated beam. In addition, because there are so many degrees of freedom, it is difficult to perform an independent dose or MU calculation for IMRT in the traditional way using tabulated measured dosimetric data.

Since IMRT has emerged, only a limited number of papers have been published on independent dose or MU calculations for this advanced delivery technique. Boyer et al. (1999) have started to investigate some theoretical aspects of MU calculation in IMRT. Kung et al. (2000) described an algorithm and its application for dynamic MLC delivery and compared the results with the dose calculated by the TPS. In their approach, the MU and leaf sequences for a patient treatment plan were used as input parameters to compute the dose at a given point with a modified Clarkson method. This paper represents the first successful demonstration of independent dose calculation for IMRT although no benchmark tests have been described against measurements. In another early paper on independent dose calculation for IMRT Xing et al. (2000) tried to bridge the gap between the intuitive approach described by Kung et al. (2000) and the theoretical considerations of Boyer et al. (1999). They stated an acceptance level of about ± 5% compared to the TPS and measurements for the verification of the dose at a single point in the high dose region. Later this group extended their model by including head scatter for dose calculation at an arbitrary spatial point (Yang et al., 2003). The methods described by Kung et al. (2000) and Xing et al. (2000) are based on a two-step procedure. First the fluence or MU map is obtained, which is convolved with a scatter kernel
in a second step. Linthout et al. (2004) have developed and tested a method where the dose of each sub-field, having a uniform intensity, is calculated. The sub-field contributions are summed up to achieve the total dose. Their approach was based on traditional dosimetric parameters, such as output factors, depth dose parameters and off-axis ratios, and has been applied to sliding window IMRT delivery with a micro-multileaf collimator. As a result of 166 tested IMRT beams they proposed an acceptance level of ± 5% or ± 2 cGy per beam. All publications cited above deal with step-and-shoot or sliding window IMRT delivery based on a multileaf collimator. At present, these IMRT techniques can be considered as the most common ones, at least in Europe. For completeness it should be mentioned that an independent MU calculation has also been developed for axial (serial) tomotherapy IMRT delivered with the MIMiC multi-vane collimator (Chen et al., 2002). Helical tomotherapy devices are currently installed in several hospitals in Europe and their specific properties and QA aspects have been discussed in Section 4.2.1.

During the last years, software tools have become commercially available to perform independent MU calculations for IMRT. Most of the applied algorithms have either been based on the method described by Xing et al. (2000) or on the one described by Kung et al. (2000). However, all these independent MU calculators are not restricted to IMRT. They also cover irregular MLC-shaped beams, dynamic wedges or electron beams. Important common features of these commercial software tools are that both step-and-shoot and sliding window MLC delivery are supported, leaf sequence patterns including the number of monitor units can be directly imported from the TPS, and dose calculations are performed in the dose specification point located on or off the central axis in a flat homogenous semi-infinite water phantom.

Besides the advantages of an independent dose calculation method for IMRT, which is in general less time consuming than experimental methods for patient-specific QA, there are some disadvantages too. Firstly, independent MU calculation tools for IMRT require extensive verification of the software, i.e. they must be commissioned prior to their clinical use. Secondly, because calculations are performed in a homogeneous medium, larger deviations can be expected when comparing dose values from the patient, based on CT information, and independent dose calculations performed in a flat-water phantom. Deviations due to density variations, surface curvatures or oblique beam incidence might occur. A simple option is to perform a path length correction for dose calculation in an independent IMRT dose calculation. Alternatively, IMRT plans can be recalculated in a homogenous water phantom (or any other IMRT verification phantom) and results of such a recalculated plan can then be compared with an independent IMRT dose calculation method. As an example, Figure 5.5 illustrates the results of an independent calculation of the dose at the isocentre for IMRT plans obtained for 34 prostate patients (206 fields) treated with a sliding window IMRT technique at Santa Maria Nuova Hospital, Reggio Emilia, Italy. As can be seen from Figure 5.5, performing dose calculations in a homogeneous phantom instead of using patient CT data reduces the deviations between calculations performed by the TPS and independent dose calculations for
prostate IMRT treatment fields. Results of single field measurements are also shown giving on average the same result with a somewhat smaller standard deviation, which might be due to the limitations of the independent dose calculation algorithm. Obviously it takes experience and proper commissioning to know the cause of these discrepancies.

Figure 5.5 Deviations between independent dose calculations, calculations performed by the TPS and measurements for single fields of prostate IMRT treatments.

ESTRO has also initiated a project on MU verification for advanced treatment techniques. The main difference of this project and existing independent dose calculation methods was the intention to use an energy fluence-based approach and a pencil beam model rather than traditional empirical models. The head scatter model has been described by Olofsson et al. (2003, 2006). It considers primary energy fluence and scattered radiation from an extra focal source and from secondary collimators, as well as backscatter to the monitor chamber as a function of collimator setting and treatment head design. This general head scatter model was verified for 19 photon beam qualities in the range from 4 up to 50 MV, provided by nine different treatment units from six manufacturers. The pencil beam model uses parameters, which can be derived from the quality index of a high-energy photon beam (Nyholm et al., 2006). Thus the main advantage of this energy fluence-based approach is that it requires only a few and easily obtainable input data to tune the semi-analytical model. Moreover, the whole algorithm is more versatile and powerful. Similar to other commercial independent MU programs, irregular beams, physical and dynamic wedges and dose points located off the collimator axis can be handled, while file transport from the planning system to the MU
verification (MUV) software is enabled directly through a Dicom interface. Recently, a beta version of MUV was distributed to test the model in different reference centres for conformal radiotherapy as well as for IMRT (Georg et al., 2007a,b). In Section 6.2.6 the application of MUV for IMRT verification in the Department of Radiotherapy of the Medical University Vienna will be described. Another interesting and novel feature is the integrated error estimation. From benchmark tests against measurements an error model was established. Thus the final result of an independent dose calculation includes an error estimation based on known uncertainties of the model. The final aim is to draft an ESTRO booklet, which describes the model. Results obtained from testing this independent algorithm against commercial treatment planning systems and measurements have recently been published (Georg et al., 2007a). ESTRO-EQUAL will release a CE/FDA certified software package called EQUAL-Dose for independent dose calculation with the semi-analytical dose calculation model that is based on the MUV software described in this chapter.

For an independent dose calculation for individual IMRT beams or a composite IMRT treatment plan it must be realised that in principle a completely and independent software solution is necessary, either a commercial or an in-house developed one. Furthermore, in some of these systems body surface curvature and tissue inhomogeneities are not taken into account, which may limit their application. Also secondary effects may not be modelled as accurate as in the clinically applied software and may influence the final results of such a type of IMRT dose verification. For example, leaf transmission, rounded leaf ends, and the tongue-and-groove effect have an influence on the result of independent IMRT dose calculation, at least for some fields and dose calculation points. At present, these effects are well known and understood but it is not clear to which extent these effects should be taken into account or how accurate they need to be modelled.

The expectations on independent dose calculation methods for patient-specific QA in IMRT are high. On the other hand it needs to be recognised that independent MU calculation for IMRT cannot replace experimental methods for commissioning IMRT equipment. Also it should be realized that most of the independent dose calculations verify only the dose at a single point, which is a significant limitation of these methods compared to, for instance, a recalculation of an entire treatment plan through a patient with Monte Carlo. The commercial interest in independent MU calculation for IMRT is growing and tools have become commercially available. Although the workload for patient-specific QA of IMRT can often be reduced by using an independent calculation procedure, it should be noted that such an approach does not include a check of the correct file transfer from the treatment planning system to the treatment delivery console, or a verification of the MLC performance for IMRT delivery. Independent MU calculation is just another dose calculation (barrier or safety net) that may reduce or reveal errors/inconsistencies in the treatment planning system/process. Clinical application of independent dose calculation tools for IMRT verification just started on a larger scale and more experience is needed until final conclusions can be drawn on its place in a QA programme of IMRT.
6. PATIENT-SPECIFIC QA PROCEDURES

Ideally patient-specific QA procedures should be performed during the actual patient treatment, i.e. by using in vivo dosimetry methods, preferably applied by the radiation therapy technologists (radiographers). These methods are, however, still scarcely employed and most QA procedures are performed prior to treatment. Pre-treatment patient-specific QA procedures can be divided into field-by-field and total plan verification, as discussed in Chapter 3. Field-by-field verification has the advantage that errors are directly linked to a specific field or segment, and the source of error can easily be traced. The impact of these errors on the actual patient treatment is, however, difficult to predict and therefore verification of the total (composite) plan is applied by many groups as discussed in Chapter 1 and elucidated further in Chapter 3 using the “conceptual pyramid” approach. Possible differences in the individual fields are now expressed in terms of the total dose distribution, thus allowing estimates of their clinical impact. Also field-by-field verification is often performed under zero-degree gantry angle thus missing effects of incorrect gantry angle or gravity on the MLC position during dose delivery.

Patient-specific pre-treatment verification should be kept as simple as possible because the QA efforts are proportional to the number of patients. On the other hand they should be extensive enough to be able to detect errors and problems that may occur with the specific combination of TPS, sequencer and delivery equipment. This is usually a sensitive balance depending on specific local circumstances such as accuracy required and resources available. For instance, in the ease of using gamma evaluation to set criteria for agreement, quantitative information is lost. The main goal should be to detect gross errors. A reasonable approach is therefore to perform experimental dose verification or independent dose calculations with action levels that may be relaxed. As a matter of fact, it can be argued that more effort should be put into the QA of equipment used for IMRT and the QA of class solutions as discussed in Section 3.5. It is the main purpose of this chapter to give examples of practical solutions of patient-specific QA procedures applied routinely in a number of European centres.

6.1 PRE-TREATMENT VERIFICATION

6.1.1 Ghent University Hospital, Ghent, Belgium

A practical patient-specific QC test is to deliver the actual patient treatment plan to a representative phantom that contains one or more point detectors. Figure 6.1 displays the phantom used for QA of prostate IMRT patients, having an ionisation chamber mounted at the isocentre, as routinely applied at Ghent University Hospital (GUH), Ghent, Belgium. The method assumes that the reading of the ionisation chamber for the number of MUs applied clinically is proportional to the weight of the patient. Furthermore it is necessary that the isocentre lies...
in a uniform high-dose region, a condition satisfied for prostate IMRT treatments. Figure 6.2 represents the difference between measured and planned dose, using the patient CT data for a contiguous series of 236 prostate patients treated at GUH. From these 236 patients, 185 were treated with a 3-field technique, 1 with 5 fields, 3 with 6 fields, 25 with 7 fields and 22 patients were irradiated with an 8-field technique. Each beam of the 3-field treatments had a weight of approximately 90 MUs. The number of MUs of the other beam configurations varied. For planning purposes use is made of a homemade optimisation programme (De Gersem et al., 2001) in combination with a commercial TPS (Philips, Pinnacle) for the final dose calculation. Because the same phantom is used for all patient verifications, and heavier patients need in general more MUs than thinner patients, higher readings are observed for larger patient weights. A regression analysis depicted that most of the difference can be explained by the differences in patient’s weight only. If the residual difference was larger than 5%, the MU verification measurement was repeated. If the difference remained larger than 5%, the dose distribution was computed using the CT data of the reference phantom instead of using those of the patient. In 15 of the 17 cases that exceeded the 5% tolerance level, the recomputed dose was found to be within 5% of the measured dose. This means that for those 15 patients the difference between the phantom and the actual patient geometry was the reason for the observed difference. Only in two of the 236 cases, an error in the treatment plan was the origin of the observed difference. In both cases a premature plan rather than the optimised plan was erroneously sent to, and executed by, the accelerator.

![Reference phantom set-up for experimental MU verification of prostate IMRT treatments at Ghent University Hospital, Belgium. The number of MUs of the unmodified patient treatment plan is delivered to the phantom that contains an ionisation chamber at the isocentre.](image)
Figure 6.2  Linear regression analysis of the dose difference between measured and planned dose, using the patient CT data, at the isocentre versus patient weight. The resulting coefficient of correlation is 0.78. Only 17 of the 236 patients (7%) have dose differences that deviate by more than 5% from the regression line.

6.1.2  Santa Maria Nuova Hospital, Reggio Emilia, Italy

At Santa Maria Nuova Hospital in Reggio Emilia, verification of IMRT treatments of head-and-neck patients, generally given with 7 fields applying the sliding window technique, is performed by a field-by-field verification using a 2D detector array (PTW Seven 29, see Section 3.3) embedded in a water-equivalent phantom in such a way that the points of measurement are at 5.0 cm depth. To improve the spatial resolution four measurements are performed by shifting the device 5 mm in the X- and Y-direction and combining them via the PTW software. Gamma evaluation is then applied with a dose-difference criterion of 3% and a DTA criterion of 3 mm. A plan is accepted if the percentage of points with gamma below one is higher than 90%. In Figure 6.3 the results are shown for the verification of 130 head-and-neck IMRT fields. These cases are planned using the Varian Eclipse TPS with a maximum number of 80–100 segments, although for cases having steep fluence gradients higher values, up to 120 segments, are used. The mean value of the total number of MUs is about 1400. Almost all patients are treated with an integrated simultaneous boost having fraction sizes of 1.8, 2.0, or 2.2 Gy for the elective nodes, regional disease and primary target volume, respectively.
Plan verification of prostate IMRT treatments is performed in this centre by means of an independent MU calculation as elucidated in Section 5.2. A plan is accepted if the total dose at the isocentre calculated with the independent MU programme does not differ more than 3% from the dose calculated with the TPS.

### 6.1.3 Tübingen University Hospital, Tübingen, Germany

At Tübingen University Hospital all IMRT treatments are optimised with Monte Carlo dose computation having already, in principle, a high accuracy as outlined in Section 5.1. For that reason only a verification of the correct transfer of the plan from the TPS to the linac is performed, in combination with an intensive programme verifying the accurate performance of the MLC. In addition, the routine patient-specific QA programme includes a verification of the total number of segments and MUs of the treatment plan transferred to the linac. This “plausibility check” relies on the fact that the total number of photons irradiating the target volume must be equivalent for IMRT and 3DCRT using conformal open (non-wedged) beams. Intensity modulation is merely a redistribution of photon fluence within each beam and between beams. Following this argument, the quantity mean fluence (in MUs) = sum (segment area x MU) / (beam area), summed up over all beams, should result in a number that is comparable to the number of MUs required for a 3DCRT treatment of the same patient and a similar distribution of beam directions.
The procedure is illustrated using the case of a 4-year old girl presenting after surgery and chemotherapy with an embryonic rhabdomyosarcoma in the abdomen. Irradiation of the whole abdomen with 17 fractions of 1.6 Gy was intended to control microscopic disease and possibly disseminated tumour cells. The goal was to keep the dose in the kidneys and liver as low as possible while at the same time to irradiate the entire circumference of the liver. The treatment plan was generated with the in-house developed IMRT treatment planning system Hyperion and consisted of 12 coplanar 6 MV beams. The system uses Monte Carlo dose computation during the final optimisation of MLC segment shapes and weights, so that the final dose distribution is identical with the result of the optimisation. Figure 6.4 shows the resulting plan.

![Figure 6.4](image)

Figure 6.4 Whole abdomen IMRT treatment of a 4-year old patient with a rhabdomyosarcoma with sparing of liver, kidneys and spinal cord.

In this case, 281 beam segments and 1721 monitor units (minimum number of 3 MUs per segment) were needed for the total plan and the mean fluence was 96.1 MUs for all twelve beams. Given the small diameter of the patient, 96.1 MUs for 1.6 Gy per fraction appears reasonable. (Reference conditions are 100 MUs = 1 Gy at 10 cm depth, SAD = 100 cm).

Normally, IMRT dose distributions are not verified experimentally and are only checked for plausibility as described above. For this particular case, a full verification with film measurements was also performed for each beam with vertical incidence on a cubic phantom. Measurements were performed at 3, 10 and 20 cm depth. One of these verification films is shown in Figure 5.4. Overall, the agreement was almost everywhere within the 2%/2mm acceptance threshold in the majority of the points. This high level of accuracy requires a meticulously tuned accelerator head model in the Monte Carlo calculations. Some effort has therefore to be invested into the robust and efficient determination of head model parameters such as primary source distribution and energy spectrum, as explained in Section 5.1.
IMRT treatment verification is done at the Seville University Virgen Macarena Hospital by performing an independent dose calculation using an in-house developed Monte Carlo simulation system (Leal et al., 2003). For that purpose a cluster of 140 Pentium III 700 MHz and Pentium IV 2.4 GHz PCs with 256 Mbyte memory is used in the department of the University of Seville. As an example the result of a MC verification of the dose distribution obtained with the in-house forward IMRT planning system for a treatment of a nasopharynx case is shown in Figures 6.5 and 6.6. A 6-field technique was used having 13 segments each, which required 497 MUs. As can be seen from Figure 6.5, the largest differences are found near and in the air regions along the larynx where the dose calculation performed by the clinical TPS has some limitations and the MC calculations give more accurate results. Acceptability of a plan is based on a gamma evaluation, shown at the bottom part of Figure 6.5, using as action level that no more than 5% of the points should have a gamma value larger than one. Because not all dose distributions failing to pass this action level may result in a clinical unacceptable plan, a clinical examination is also performed. In this particular case the differences occurred mainly in the air cavities and therefore do not have clinical consequences. Furthermore, such an examination also considers tolerance levels for each organ, and demands an acceptable DVH for both dose calculation methods. In Figure 6.6 the DVHs obtained by the MC and TPS calculation are given. Despite the differences observed from the gamma evaluation, the

**Figure 6.5** Verification of an IMRT treatment of a nasopharynx tumour showing in the upper part the isodose lines generated by the TPS (blue) and the Monte Carlo simulation (red). The bottom part represents a gamma evaluation applying 3%/3mm criteria.
final clinical evaluation considered this plan acceptable because the MC calculations showed a lower dose and a more accurate result for the OARs, especially for the parotids.

![Figure 6.6](image)

**Figure 6.6** Dose-volume histograms of the PTV and organs at risk resulting from the TPS dose calculation and the Monte Carlo simulation of the nasopharynx case presented in Figure 6.5.

### 6.1.5 University Hospital U.C.L.-St. Luc, Brussels, Belgium

In St-Luc University Hospital, patient-specific QA is performed for each helical tomotherapy treatment using the “cheese phantom” as elucidated in Section 4.2.1 (see Figure 4.8). Results of a comparison between ionisation chamber measurements and TPS dose calculations are shown in Figure 6.7. If deviations larger than 4% occur, visible inspection of the position of the ionisation chamber with respect to the dose distribution is performed, generally showing that the ionisation chamber was partly located in a high dose gradient region. Gamma evaluation of the film measurements is applied with a dose-difference criterion of 3% and a DTA criterion of 3 mm. Figure 6.8 shows an example of a verification of an IMRT treatment of an oropharynx tumour using a film measurement. As elucidated in Section 4.2.1 tomotherapy does not work on a monitor unit-based system. For that reason no information about MUs is given in this example.
Figure 6.7 Results of 185 ionisation chamber measurements yielding a mean deviation of $-0.5\%$ and a standard deviation of $1.9\%$.

Figure 6.8 Verification by means of film dosimetry of a helical tomotherapy treatment of an oropharynx case showing the isodose lines generated by the TPS (solid lines) and the measurements (dotted lines).

6.1.6 Medical University Vienna-AKH Wien, Vienna, Austria

At the Department of Radiotherapy of the Medical University Vienna, verification of step-and-shoot IMRT treatments is done by performing an independent dose calculation since 2006. For that purpose the software developed in the ESTRO project (see Section 5.2) is
The procedure consists in a recalculation of a patient IMRT plan for a homogeneous head-and-neck or pelvis phantom and comparing the resulting dose at multiple points (at arbitrary locations) or along a line with the dose in corresponding points determined with the clinically applied TPS (Nucletron, Oncentra Masterplan). As an example, Figure 6.9 shows the results of an intercomparison of the monitor unit verification (MUV) data along a line in the composite plan of a head-and-neck patient with both film measurements and the dose calculations performed by the TPS using two different algorithms. The IMRT technique represented in this figure consisted of 7 beams (10 MV), 69 segments and 513 MUs in total. IMRT planning is generally performed using 2 MUs, 4-cm² field size and 2 open leaf pairs per segment as a minimum, and 15 segments and 10 intensity levels per beam as a maximum. From the results shown in Figure 6.9 it can be concluded that the physics model underlying the MUV software is very accurate for these types of IMRT plans.

Before applying the software clinically, the dose calculated in the high dose and medium to low dose regions, using the MUV software, was compared with measurements and calculations performed by the TPS (Georg et al., 2007a,b). Ionisation chamber measurements, using a calibrated 0.3 cm³ Farmer-type chamber, and film dosimetry were used to check both composite IM treatment plans and individual beams. For a total of 52 composite IMRT plans (mainly in the head-and-neck and pelvic region) the mean deviation between the MUV calculations of the dose at the isocentre and the ionisation chamber measurements was 1.1 ± 2.9% (1SD) with a maximum deviation of 14%. The results of the verification of these 367 individual IM beams are shown in Figure 6.10 demonstrating that only in a few cases large relative discrepancies were observed. However, absolute dose deviations per beam were mostly smaller than 3 cGy, with a mean deviation of only 0.3 ± 0.8 cGy.

Based on results obtained in these benchmarking studies of MUV, for full IMRT plans and points close to the isocentre a confidence limit of 3% dose deviation (with respect to the
prescribed dose) or 6 cGy are applied as action levels in clinical routine. As long as points of interest close to the isocentre are considered, the same criteria are used for dose points in the target volume or in organs at risk. For off-axis points at distances larger than 5 cm or for points in low dose regions a 5% dose deviation (with respect to the prescribed dose) or 10 cGy are considered as acceptable values. In cases where there are larger deviations, ionisation chamber measurements and film dosimetry have to be performed. Depending on the outcome of such further investigations, the number of MUs will be rescaled if there is the “trend” of a general offset in the dose distribution. If the dose distribution cannot be reproduced, the plan is rejected and re-planning has to be performed.

Figure 6.10 Relative deviation in [%] between dose calculations performed with the treatment planning system and the independent dose calculation software MUV for individual IM beams. All deviations displayed refer to a single calculation point per plan that was located in the highest dose region. The absolute deviation (with respect to the prescribed dose) was generally less than 3 cGy; only in 4 cases deviations between 3.5 and 4.0 cGy were observed.

Such a MU verification programme reduces the workload considerably compared to, for instance, film measurements for each patient. However, the influence of patient anatomy is neglected by using a homogeneous, simple phantom. Such a procedure may therefore introduce a small deviation from the “true” dose distribution, as elucidated in Section 5.2. Also the verification of the data transfer from the TPS to the treatment unit is not included in a verification procedure using the MUV software. Therefore, a dry-run is performed to verify that the correct information is transferred to the linear accelerator for delivery of the IMRT plan.
6.1.7 German Cancer Research Center, Heidelberg, Germany

At the German Cancer Research Center, DKFZ, patient-specific IMRT QA is divided into two parts; during the first part a pre-treatment absolute dose verification is performed, while the second part consists of a fluence verification in front of the patient during the first irradiation fraction.

**Part 1:** A fast dosimetric total plan verification procedure was implemented by measuring the absolute dose with 5 ionisation chambers (0.125 cm³) inside a solid water matrix ionisation chamber phantom (Figure 6.11) located at positions in a high dose region having a low dose gradient. The mutual interference of the ionisation chambers was investigated and found to be negligible (Rhein and Häring, 2005). The procedure is fast: five dosimetric verifications can be done within 1 hour. This part of the patient-specific QA programme is done once a week for all new IMRT patients. The right part of Figure 6.11 shows the frequency distribution of the percentage deviation between dose computations and ionisation chamber measurements for 262 mainly head-and-neck IMRT cases and 1131 measurement positions inside the matrix phantom. The average deviation is 0.0% with a standard deviation of 2.1%. The tolerance level for plan acceptance was defined to be to ± 5% for a single measurement point and ± 3% for the average of the total number of measurement points to be compared for one IMRT case.

![Figure 6.11](image)

**Figure 6.11** Left: Solid water matrix ionisation chamber phantom. Right: Frequency distribution of percentage deviation between ionisation chamber measurements and dose computation (262 IMRT cases and 1131 measurement points).

**Part 2:** During the first fraction of an IMRT patient treatment a special film tray is inserted into the accessory holder of the treatment machine. A radiographic film is irradiated to measure the entrance dose distribution for each IMRT field. The design of the film tray in-
cludes guidance needles to mark a co-ordinate system for correlation, and a copper plate as build-up material for the film and to eliminate electron contamination of the beam (Figure 6.12). The primary fluence histograms resulting from the inverse treatment planning system are first converted to smooth fluence distributions using a convolution kernel based scatter model, next corrected for MLC leakage, and then compared with the measured IMRT intensity distribution. The evaluation of the two fluence matrices includes a fluence level and leaf position comparison using a 5% fluence maximum and 1 mm distance-to-agreement tolerance level. Fluence profiles, a fluence difference matrix and a gamma evaluation are included in the software tool (Figure 6.12).

Figure 6.12 Top left: Film-tray for fluence measurements. Top right: Gamma evaluation of an IMRT field using a 5%/1mm tolerance level. Bottom: Comparison of measured and calculated intensity profiles for the same field as shown in the top right figure.
6.1.8 San Raffaele Hospital, Milan, Italy

At San Raffaele Hospital, patient-specific QA is performed for each helical tomotherapy treatment using the “cheese phantom” described in Section 4.2.1. Routinely, ionisation chamber measurements are performed at 2 points (10 mm under and 5 mm above the film) in the sagittal plane while the film is placed in the coronal plane, as shown in Figure 4.8. The film dose distribution is normalised relative to the average between the two dose values measured with the ionisation chamber. Comparison between TPS dose calculation and dosimetric verification is performed by means of ionisation chamber measurements for two positions (inside the target) and a semi-quantitative gamma evaluation of the 2D film measurement (chosen criteria: dose difference equal to 3% and a DTA equal to 3 mm). The results of 215 ionisation chamber measurements are presented in Figure 6.13. The mean difference between expected and measured dose was -0.4% (1SD = 2.2%). If the three anatomical regions are separately considered the mean differences were equal to 0.3%, -2.1% and 0.4% for prostate, head-and-neck and abdomen treatments, respectively, each having a standard deviation of 2.0%. These results suggest an inverse relationship between the agreement of the measured and expected dose and the complexity of the treatment; head-and-neck treatments involve more complex intensity modulation compared to prostate and abdomen treatments. When deviations larger than 4.0% occur, a visible inspection of the position of the ionisation chamber with respect to the dose distribution was performed, showing that in many cases the detector position was partly located in a high dose gradient region. Analysis of the film measurements showed that the gamma value was <1 for 90% of the data and only 2.8% of the gamma values were in the range 1.5 to 2. As elucidated in Section 4.2.1 tomotherapy does not work on a monitor unit-based system. For that reason no information about MUs is given in this example.

![Figure 6.13](image)

**Figure 6.13** Results of 215 ionisation chamber measurements of 108 helical tomotherapy treatments for three anatomical regions.
At Lund University Hospital, pre-treatment IMRT verification is performed for each patient, being mainly head-and-neck patients. A typical IMRT step-and-shoot plan consists of 7 fields with 10-15 segments per field adding up to a total of approximately 500-600 MUs per fraction. The dose plan is verified field-by-field using a 2D diode detector array (Mapcheck, Sun Nuclear, see Section 3.3). To take into account possible gravitational effects on the MLC leaf position, the array is mounted on an in-house built detector holder (Figure 6.14) attached to the accessory ring. In this way the original clinical plan remains unchanged with regard to collimator and gantry rotation. Build-up material is added to measure at 5 cm radiological depth at SDD=95 cm in a plane through the isocentre perpendicular to the central axis of the beam. Comparison between the TPS dose matrix and the detector readings is performed using gamma evaluation. The verification plan is passed if the analysis of each field shows gamma values $\leq 1$ for at least 90% of the detectors. Gamma criteria are set to 3% dose difference and 3 mm DTA based on internal analysis of a number of sources of uncertainties including beam stability, and dose calculation and set-up accuracy. Points that fail the gamma evaluation are analysed separately for clinical relevance. These points are typically found in low dose regions, such as the spinal cord where it was found that the dose calculated by the TPS (Nucletron, Oncentra Masterplan) is underestimated. This is due to inaccurate modelling of the transmission through the MLC and jaws. If the measurements show a larger area of points that fail the gamma evaluation, the error in the dose is calculated and related to the patient dose. If unacceptable risk of over-dosage is found, the treatment plan is re-optimised, generally by using larger segments.

Figure 6.14  Diode array (arrow) holder allowing field-by-field measurements with full gantry rotation.
6.1.10 University Medical Center Hamburg-Eppendorf, Hamburg, Germany

IMRT treatment verification is performed at the University Medical Center Hamburg-Eppendorf prior to each new IMRT patient treatment. It is based on a verification of the dose distribution of the total plan in combination with an independent MU calculation. Details of the applied IMRT step-and-shoot techniques are for head-and-neck treatments: 7 fields having in total 90 to 100 segments and about 600 MUs, and for prostate treatments: 5 fields having in total 40 to 50 segments and about 400 MUs. The number of intensity levels is 7-10 and the resolution of the fluence matrix 5 mm x 10 mm or 10 mm x 10 mm.

Total plan verification

For the dose verification of the total plan a cubic phantom of 18 cm each side made of RW3 solid water (Gammex/RMI, Middleton, WI, USA) was developed (see Figure 6.15). It is labelled as EasyCube and commercially available (Euromechanics, Schwarzenbruck, Germany). The phantom is used with EBT radiochromic films (International Specialty Products, Wayne, NJ, USA) and ionisation chambers. The calibration of the film is performed with a 3 x 3 field pattern of 2 cm x 2 cm field segments of different dose values, having a one-cm gap between the segments (see Figure 6.15). The segments cover a dose range of 0.1 to 3 Gy. The calibration has to be repeated for each lot of new films. (See also Section 3.2). For the total plan evaluation three to five films are positioned in parallel planes in the phantom. In principle the phantom can be used in the transversal, sagittal and coronal direction. For routine QA checks it is mainly used in the coronal direction. For a total plan verification a new treatment plan is generated in the TPS (CMS, XiO) by transferring all treatment parameters (e.g., beam angle, segments, MUs) to the EasyCube. The position of the phantom for this hybrid plan should cover the region of interest for verification. If possible the centre of the EasyCube should be identical to the isocentre, but any other position can be used as well.

Figure 6.15 The EasyCube phantom (left) and the calibration film (right) used for the verification of the hybrid plan.
The irradiated films are scanned by using a commercial document scanner with transparency capability (Microtek ScanMaker 8700 or an Epson Expression 10000XL). From the TIFF-files only the red channel is used for further evaluation. The scanner signal is converted into dose using the calibration film. A comparison between the measured and corresponding reference dose distribution from the hybrid plan is performed as dose difference in terms of Gy. Areas with dose differences larger than 0.2 Gy should be below 5% of the irradiated area. If this criterion is not met several options are possible: the scanning direction of the film is checked, the film is rescanned, rescanning and re-evaluation is performed, a new verification plane is chosen or a new calculation of the plan with a modification of the grid size is performed. If none of these options results in an acceptable plan, a new treatment plan is generated.

The total dose is measured with a PinPoint ionisation chamber having a 0.015 cm³ volume (See Section 3.1). Different plugs for positioning the chamber in the EasyCube are available thus allowing its positioning in a grid of 1 cm width. If the difference between measured and planned dose is smaller than 3%, the plan is accepted. For differences between 3% and 5% another point in a more homogeneous dose area is chosen, which resulted in most cases in an acceptable plan.

*Independent MU calculation*

To verify the number of monitor units calculated by the TPS, an in-house developed program is used to calculate the dose contribution from each field at one point of interest (POI). The algorithm is based on the calculation of the primary radiation and the superposition of scattered dose for all MLC segments of each field. The convolution kernel is derived from TMR measurements of square field sizes in water. The program uses the treatment parameters of the separate fields (e.g., number of MUs and shape of each segment) from the data transfer file that delivers planned data to the linac database (RTP or DICOM files) for calculation of the dose at the POI. As the calculation is done for the phantom, the calculation depth of each field can easily be obtained from the transferred data set. If the deviation of the MU check is smaller than 5 % and the deviation between measured and planned dose at the POI is below 3 %, the plan will be accepted.

### 6.2 IN VIVO DOSIMETRY

Pre-treatment verification using phantoms is a useful method to trace errors both in the treatment planning and the delivery process. It does, however, not give information about the dose delivered to the patient and in vivo dosimetry is required to verify the dose delivery during the actual patient treatment. For instance, pre-treatment verification may result in an acceptable plan, but if the patient moves just before or during treatment the dose may be delivered at the wrong place inside the patient. Also the patient geometry may differ during
the treatment from that on which the planned dose calculation is based. Many methods of \textit{in vivo} dosimetry are available, but the use of diodes is probably still the most popular system. The dosimetric characteristics of diodes and the experience with their clinical use have been summarized in several reports (\textit{e.g.}, Huyskens \textit{et al.}, 2001, AAPM 2005). Although their use for the verification of conventional conformal techniques has proven to be very useful, the application of diodes in IMRT beams is generally impossible because of the steep dose gradients in IMRT fields. For that reason other \textit{in vivo} dosimetry approaches have been explored to verify treatment delivery during IMRT as will be discussed in the next sections.

\textit{In vivo} dosimetry can be done in two ways: a suitable probe is inserted into the patient to measure the dose directly at the required position but invasively, or a detector is placed against or at some distance from the patient to measure the dose non-invasively. In the latter case the dose inside the patient is calculated with a suitable mathematical model, using the dose measured at, or close to, the position of the detector. TLD and MOSFET detectors are often used to measure the dose at the intended target site, while the use of dosimeters positioned on the skin of a patient, as well as film and EPID dosimetry, belong to the second category of \textit{in vivo} dosimetry. In the next sections a few examples will be given of the use of TLDs, MOSFETs and EPIDs during the clinical introduction of IMRT in several centres. In Section 4.2.1 the use of radiographic film for \textit{in vivo} dosimetry during helical tomotherapy has been described. It should be noted, however, that the routine clinical application of \textit{in vivo} dosimetry for IMRT verification is still limited.

\subsection*{6.2.1 Copenhagen University Hospital, Copenhagen, Denmark}

At Copenhagen University Hospital head-and-neck patients are treated using 5 to 7 fields at equidistant angles, using the sliding window technique. Each beam is divided into two parts, caused by the relatively small maximum field size of the MLC construction. Each beam has about 150-200 segments, as calculated with the TPS (Varian, Eclipse) and generally 350-400 MUs/Gy are given. In Copenhagen \textit{in vivo} dose verification of IMRT of head-and-neck cancer patients has been applied at the introduction of IMRT and has evolved through several stages (Van Esch \textit{et al.}, 2002, Engström \textit{et al.}, 2005). In the initial phase of IMRT implementation, dose verification was carried out using intracavitary TL dosimetry. A flexible suction catheter was used as a naso-oesophageal probe containing LiF rods interspaced with small lead markers. At the first day of treatment the probe was inserted, deep enough to extend across the entire treatment volume as shown in Figure 6.16. Immediately after treatment, orthogonal images were acquired at the simulator to verify the position of each TLD. Corresponding point dose values were then extracted from the TPS and compared with measurements. Analysis of early TLD readings showed an average measured to planned dose ratio of $1.002 \pm 0.051$ (1SD, N=177). In Figure 6.16, TLD reading vs. position has been plotted for a typical nasopharynx patient. The tolerance criteria were set at 5 mm distance-to-agreement or
5% dose difference at the central point. At a later stage, the simulator images were replaced by orthogonal EPID images, which simplified the procedure and eliminated the risk of probe movements caused by patient relocation to the simulator. Since especially the distal part of the oesophagus can show a large lateral flexibility, simply scanning the patient with the probe inserted will not suffice to determine the position of and dose in the detector.

As the initial measurements showed very good agreement with the planned data, but also because the patient flow increased significantly, the cumbersome TL dosimetry procedure was replaced by the use of the flexible IDF-thin diode detector (SCX-Wellhöfer, Schwarzenbeck, Germany) inserted into the catheter. Thus a similar probe was used, having one point of measurement. The diode is positioned in the high dose region together with a tungsten pellet to mark the position. Data collected from 60 patients measured with the diode probe showed a measured to planned ratio of $0.995 \pm 0.025$ (1 SD). A less invasive method was applied by placing the detector in the mouth, between the teeth and cheek. However, the placement of the detector had a large uncertainty and was often at the edge of the treated volume, causing sometimes large deviations.

Because the results were generally inside the 90 to 110% range, with some outliers due to positional uncertainties, it was decided to stop with in vivo dosimetry and only use an independent MU calculation for dose verification at one point, applying the MUV software elucidated in Section 5.2. A deviation of 3% for the total dose per fraction and 5% per beam are the clinically applied action levels when using MUV. If the deviations are larger, a new plan

**Figure 6.16** Left: Photograph and EPID image of the oesophageal tube containing ten encapsulated TLD rods (white crosses) interspaced with lead markers. The black arrow marks the isocentre. Right: TLD readings of two measurements performed on a nasopharynx patient with one-week interval. The solid lines represent upper and lower tolerance levels (from Engström et al., 2005).
is made without tissue inhomogeneity corrections, the MUV software does not yet include inhomogeneity corrections, or the dose is calculated at other points in the target volume. If the plan is still not acceptable, a 2D verification is performed using EPID dosimetry to assess the area of disagreement for clinical evaluation.

### 6.2.2 Centre Antoine-Lacassagne, Nice, France

_In vivo_ dose measurements were performed in Nice, France, with metal oxide semiconductor field effect transistor (MOSFET) dosimeters during IMRT of oropharynx and nasopharynx tumours (Marcie _et al._, 2005). Using tissue equivalent material to take impressions of the teeth, a dentist prepared an oral plate moulded to the shape of the upper palate that contained two catheters. Inside these catheters a MOSFET as well as two lead seeds were placed, allowing localization of the detector in the oral cavity on CT slices and on verification films (See Figure 6.17). The co-ordinates of the MOSFETs were used as input in the TPS and the calculated dose values were compared with the measured data for 48 treatment plans for 21 patients. Dose determination at the point of measurement was difficult because this point was often located in an area with a steep dose gradient. Inter- and intra-fraction patient movement, determined with images of the lead seeds on film, yielded deviations up to 6.5 mm in the anterior-posterior direction, resulting in an uncertainty of 5% to 10% in the dose determination. Despite these variations in patient position, the overall result showed agreement of the mean value of the measurements with the calculated dose values within 2%.

![Figure 6.17 Left: Molded mouth plate with the two catheters. Right: CT scan showing the position of the seeds and the planned dose distribution (from Marcie _et al._, 2005).](image)

### 6.2.3 Netherlands Cancer Institute, Amsterdam, The Netherlands

At the Netherlands Cancer Institute a back-projection algorithm has been implemented clinically to reconstruct the dose in 3D within a patient based on portal images recorded with an amorphous silicon type of EPID during all IMRT treatments. The back-projection algorithm
includes a parameterized description of EPID scatter, scatter from patient to the EPID, and scatter and attenuation within the patient. The EPID-calculated dose for each treatment field is compared directly with the planned dose distribution in the plane in the patient intersecting the isocentre perpendicular to the beam direction. The γ-evaluation method is used to compare 2D dose distributions in a plane through the isocentre parallel to the EPID perpendicular to each beam. As an example the results for a 5-field 18 MV step-and-shoot IMRT plan, using about 40 segments in total derived from the TPS (Philips, Pinnacle), for prostate cancer treatment is shown in Figure 6.18 illustrating that generally good agreement exists between pre-treatment and in vivo EPID dosimetry results.

In vivo dosimetry is performed during the first three fractions of an IMRT treatment. Action levels are based on current clinical experience and are, for instance, for prostate treatments: 90% of the points should have a γ-value below one, the mean γ-value should be lower than 0.50, and the maximum γ-value should be lower than 2.0. In addition, the dose at the isocentre should not deviate more than 3.0 % from the prescribed dose. If one of these criteria is not met, a physicist evaluates the plan. A more detailed discussion on the use of these action levels can be found in Section 7.2. Differences between dose distributions derived from an EPID and a plan are for prostate treatments mostly related to changes of the patient anatomy during treatment (gas pockets) as present in field E in Figure 6.18.

**PRE-TREATMENT FIELDS (phantom): (A) 0° (B) 0° (C) 0° (D) 0° (E) 0°**

**IN VIVO TREATMENT FIELDS (fraction 3): (A) 260° (B) 320° (C) 0° (D) 40° (E) 100°**

**Figure 6.18** Gamma evaluation maps using back-projected EPID dosimetry: Top row pre-treatment and bottom row in vivo results (fraction 3) for a prostate treatment with a ‘cold spot’ in field D, where the EPID measured a 20% under-dosage in an area 0.5cm². Gamma criteria of 3% and 3mm were used.
However, in field D a discrepancy is observed due to an inaccurate dose calculation, related to the tongue-and groove parameter in the TPS, in a small area of overlapping segments (<1cm²) as shown in Figure 6.19 (McDermott et al., 2006b). Consequently, when the radiation passes first through the tongue part in one segment, and then through the groove part of the leaf in an abutting segment, in the junction area there will be an under-dosage of adjacent leaves, which may amount up to 25% (e.g., Webb, 2006).

Because of the high accuracy of the method, as well as the limited time it takes for performing and analyzing the measurements, EPID in vivo dosimetry is currently used for all IMRT treatments in the Netherlands Cancer Institute. The next step will be the clinical use of the 3D reconstruction model for the verification of the dose delivered to the target volume(s) as well as to organs at risk. Such an application has already been shown to give accurate results for breast cancer treatments using tangential fields (Louwe et al., 2003), but the method still has to be tested for IMRT fields.
7. GUIDELINES

7.1 ACCURACY IN CONVENTIONAL RADIOTHERAPY AND IMRT

From the information summarised in several papers and reports (e.g., Brahme, 1988, Mijnheer et al., 1987, Mayles et al., 1999) one may conclude that a dose deviation of 5% to 7% can be clinically detected in a number of situations. Transforming this observation into an accuracy requirement yields that the intended dose distribution should be delivered with an uncertainty of about 3%, one standard deviation in the dose delivered to the target volume. Such a recommendation resulted from data obtained with conventional or 3DCRT treatment techniques, i.e. based on the assumption that the relevant target volume was covered with a homogeneous dose distribution. Nowadays many treatments are performed using the simultaneous integrated boost (SIB) technique, often applying IMRT to obtain an optimal dose distribution, i.e. irradiating a few PTVs requiring different dose levels, including transition zones having an inhomogeneous dose distribution. For these types of SIB treatments it is not straightforward which accuracy is required for the various PTVs. Before discussing the specific aspects related to the accuracy that can, or should, be achieved for IMRT, in this section a short review is given of the clinical and technical aspects determining dose accuracy in conventional and 3DCRT.

Table 7.1 Results from studies of the accuracy of dose determinations under reference conditions for high-energy photon beams.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Region</th>
<th>Number of beams</th>
<th>Average</th>
<th>SD (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Johansson et al., 1982</td>
<td>Scandinavia</td>
<td>50</td>
<td>1.017</td>
<td>2.3</td>
</tr>
<tr>
<td>Johansson et al., 1986</td>
<td>Europe</td>
<td>16</td>
<td>1.024</td>
<td>3.3</td>
</tr>
<tr>
<td>Wittkämper et al., 1987</td>
<td>Netherlands</td>
<td>40</td>
<td>1.008</td>
<td>2.0</td>
</tr>
<tr>
<td>Hanson et al., 1991</td>
<td>International (mainly US)</td>
<td>740 (incl. Co-60)</td>
<td>1.008</td>
<td>1.9</td>
</tr>
<tr>
<td>Thwaites et al., 1992</td>
<td>UK</td>
<td>100</td>
<td>1.003</td>
<td>1.5</td>
</tr>
<tr>
<td>Dutreix et al., 1994</td>
<td>Europe</td>
<td>125</td>
<td>0.970</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>119a</td>
<td>0.985 a</td>
<td>2.5 a</td>
</tr>
<tr>
<td>Izewska et al., 1995</td>
<td>Poland</td>
<td>22</td>
<td>1.004</td>
<td>3.8</td>
</tr>
<tr>
<td>Nisbet et al., 1998</td>
<td>Ireland</td>
<td>13</td>
<td>1.002</td>
<td>1.2</td>
</tr>
<tr>
<td>Ferreira et al., 2001</td>
<td>Germany</td>
<td>114 (incl. Co-60)</td>
<td>0.996</td>
<td>2.1</td>
</tr>
<tr>
<td>Kroutilikova et al., 2003</td>
<td>Czech Republic</td>
<td>362 (incl. Co-60)</td>
<td>1.000</td>
<td>2.8</td>
</tr>
<tr>
<td>De Angelis et al., 2005</td>
<td>Italy</td>
<td>16</td>
<td>1.009</td>
<td>1.6</td>
</tr>
</tbody>
</table>

\(^a\) Excluding deviations > 12%.
In order to judge if a high accuracy can be obtained in clinical practice, several studies regarding dose measurements at a reference point in a phantom exist (see Table 7.1). For example, the ESTRO-EQUAL project reported that 637 of 669 (95%) of the studied photon beams were within a tolerance of ± 3% (Ferreira et al., 2000, 2001). The same reports show also that under non-reference conditions, the number of measurements that fail the tolerance criteria increases. For institutions participating already in an audit programme the average deviation is -0.5% with a standard deviation of 1.4%, while for institutions participating in the audit for the first time the variation increases to 2.6%. According to the latest dosimetry protocol presented in IAEA Report TRS-398 (IAEA, 2000), the dose at the reference point can be determined with an accuracy of 1.5% (1SD). Thus the EQUAL project shows that in practice the uncertainty is slightly higher even after considering the additional uncertainty of the measurement method (1.7% for the TLD-system used in that study). It should be noted that by considering dose comparisons over the last two decades the determination of dose has improved, which most likely can be contributed to the evolution of dosimetry protocols, as well as the introduction of these audit programmes. For institutions basing their dosimetry on IAEA Report TRS-398, an audit of 310 participants (excluding the EU, North America and Australia) showed an average of 1.006 with a variation of 2.4% (1SD), (personal communication J. Izewska, 2007). In conclusion, under reference conditions the dose at a well-defined point, such as the ICRU reference point, in conventional and 3DCRT with uniform intensity beams, can generally be delivered with a high accuracy, i.e., 1 to 2%, 1SD, when the total radiotherapy community (in Europe) is considered.

**Table 7.2** Results from studies of the accuracy of dose determinations in anthropomorphic phantoms of conventional and 3DCRT treatments.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Region</th>
<th>Site</th>
<th>No</th>
<th>Average</th>
<th>S D (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Johansson et al., 1987</td>
<td>Europe</td>
<td>Tonsil</td>
<td>19</td>
<td>1.035</td>
<td>3.2</td>
</tr>
<tr>
<td>Wittkämper et al., 1987</td>
<td>Netherlands</td>
<td>Prostate</td>
<td>18</td>
<td>1.015</td>
<td>1.5</td>
</tr>
<tr>
<td>Thwaites et al., 1992</td>
<td>UK</td>
<td>Pelvis-Homogeneous</td>
<td>62</td>
<td>1.008</td>
<td>2.7</td>
</tr>
<tr>
<td>Aird et al., 1995</td>
<td>UK</td>
<td>Head-and-neck</td>
<td>13</td>
<td>1.007</td>
<td>2.1</td>
</tr>
<tr>
<td>Kron et al., 2002</td>
<td>Australasia</td>
<td>Head-and-neck</td>
<td>19</td>
<td>1.001</td>
<td>3.5</td>
</tr>
<tr>
<td>Venables et al., 2003</td>
<td>UK</td>
<td>Breast</td>
<td>36</td>
<td>0.979</td>
<td>1.3</td>
</tr>
<tr>
<td>De Angelis et al., 2005</td>
<td>Italy</td>
<td>Pelvis</td>
<td>16</td>
<td>1.009</td>
<td>2.2</td>
</tr>
</tbody>
</table>
The accuracy at an institution in delivering the intended dose distribution should not be restricted to investigating the accuracy of the dose under reference conditions, but should include the whole radiotherapy chain. For instance inaccuracies in CT geometry, TPS performance and input data, beam set-up all contribute to the uncertainty in the “local dose” delivered to a patient. Furthermore, external audits should be used to scale the “local dose” to dose values in agreement with values derived using international dosimetry protocols. With the general introduction of 3DCRT, dose intercomparisons were not only restricted to comparing the dose in one or a few points in the target volume, but also looked at the complete 3D dose distribution including the dose in organs at risk.

For these more complex situations, several reports presenting audits or comparisons between delivered and calculated dose values have been published during the last 20 years (see Table 7.2). Details about the location of the points where the dose had been verified, as well as the way these dose intercomparisons were performed, can be found in the added references. From this table it can be concluded that the delivery of absorbed dose is on average also quite good, i.e. within 1 to 2% compared to the stated value. However, the variation can be rather large, with an SD up to 3.5%. Part of the variation must be attributed to the dosimetry method used during these audits, but still there is an inter-institutional variation. Thus, it is of utmost importance for a clinic to have control of their accuracy in the dose delivery. One should also note that within a certain institution the accuracy in “relative” dosimetry, i.e., the reproducibility in the dose delivery is usually better than these figures indicate since the offset in “absolute” reference dosimetry is not included in relative dose measurements. For instance, in the Australasian report the standard deviation for the head-and-neck study decreases from 3.5 to 2.3 % if dose values are expressed as “local dose values” and differences in absolute dose calibration are not taken into account. The same trend has also been observed for their prostate case, i.e. a reduction from 3.3 to 1.9 % (Kron et al., 2002).

At this moment only a few studies have been carried out, or are in progress, concerning audits or dose comparisons in anthropomorphic phantoms irradiated with IMRT beams. The results of these studies, showing the accuracy of dose determinations of IMRT treatments, have been summarised in Table 7.3.

The study of the ESTRO-QUASIMODO project has already been discussed in Section 3.2.1. The QUASIMODO results for the PTV are very good and comparable to those presented in Table 7.2 for the verification of conventional or 3DCRT treatment techniques using a pelvic phantom. For the OAR the mean deviation is also good but the standard deviation is larger; a maximum deviation of 5.8% in one of the participating centres was observed. It should be noted that the ESTRO-QUASIMODO data given in Table 7.3 concern the mean dose in the PTV and OAR as derived from seven films in combination with an ionisation chamber measurement at the centre of the PTV and OAR. Therefore limitations in the accuracy of the dose delivery or planning system calculations at field edges and out of field are not explicitly
specification was not always specified in this way. Nevertheless the agreement between measured and computed dose distributions was better than might be expected for the wide range of planning and delivery system combinations investigated in this project.

The results of a study of tests performed by the Radiological Physics Centre (RPC) of a special head-and-neck phantom mailed to 128 North American institutions wishing to participate in a Radiation Therapy Oncology Group (RTOG) trial, showed unexpected large deviations between planned and measured IMRT dose distributions (Molineu et al., 2005, Ibbott et al., 2006). Roughly one third (48/163) of all the irradiations of the head-and-neck phantom failed to meet the criteria: the ratio of the measured dose (as determined from TLDs) to institutions’ stated dose was expected to agree within 7%, and the distance-to-agreement in the high-dose gradient region near the OAR was expected to be not greater than 4 mm. Discrepancies in the dose delivery were attributed to a variety of reasons including the use of inappropriate data in the treatment planning systems and collimator positioning errors.

The physicists of the French-Belgian GORTEC group (Groupe d’Oncologie Radiothérapie des tumeurs de la TEte et du Cou) are in the process of performing a dose intercomparison using their dedicated head-and-neck phantom PIGG (Physics Imrt Gortec Group). The phantom is homogeneous and includes a unilateral CTV receiving a therapeutic dose, bilateral CTVs receiving a prophylactic dose, and organs at risk (spinal cord, parotid glands, oral cavity and larynx). Different types of ionisation chamber can be inserted at five points in the PTV and at two points in the spinal cord, as well as films to verify the dose distribution. The study will be performed in 25 to 30 institutions and be finished in 2008. Some preliminary data are shown in Table 7.3 indicating that the average of all dose comparisons at the seven measurement points, including those in regions of high dose gradients and in low dose areas,
is very good. A more detailed analysis will be necessary to separate the data for the PTV from those at the OARs.

ESTRO and the Organisation of European Cancer Institutes (OECI) joined to establish an initiative aimed at auditing new radiotherapy techniques. An initial goal was rotational delivery of treatment such as helical tomotherapy and intensity-modulated arc therapy (IMAT). A homogeneous solid water phantom, with fictitious volumes to mimic concave target volumes and organs at risk, was sent to participating centres to plan and irradiate with their helical tomotherapy facility. The radiochromic films and TLDs inserted in the phantom were sent back to the ESTRO-EQUAL laboratory for evaluation. Some preliminary results of this intercomparison are also shown in Table 7.3.

The ESTRO-QUASIMODO dose intercomparison was performed at an early stage of implementation of IMRT in most of the participating centres. Since that time a number of improvements in both the planning and delivery of IMRT treatments have been introduced. Also the experience in those institutions and other European centres with respect to IMRT verification, for instance those participating in the GORTEC group, has been increased considerably. Therefore more information is urgently needed about the accuracy of IMRT treatment delivery in Europe by having similar types of independent audit or intercomparison programmes. Since it is common practice in IMRT to perform patient-specific dose verification, each institution must have the knowledge to what degree they are able to plan and deliver an IMRT treatment. In Chapter 6 examples are given yielding information about the accuracy of IMRT delivery for specific techniques and equipment as applied in a number of centres in Europe.

The absolute value of the dose and its 3D distribution in the target volume and the OARs determine the success of a radiotherapy treatment, irrespective of the chosen technique. Therefore accuracy requirements for the dose delivery of IMRT should in principle not differ from those formulated for other types of radiotherapy. The total dose distribution delivered with IMRT techniques to a patient or a phantom differs from other 3DCRT techniques in the amount of radiation dose outside the target volume, which is smeared out over a larger volume. In this way healthy tissue, and particularly OARs, will have smaller regions receiving a high dose while more regions will get a low to medium dose. These dose distributions are generally characterised by steep dose gradients around the target volume, particularly in the direction of adjacent OARs. Dose delivery in IMRT should therefore be focussed on the precision and accuracy of the position of these steep dose regions assuming that the dose distribution in the target volume is homogeneous. It should be realised that due to set-up uncertainties, movement of organs and instability of treatment equipment the actual dose distribution in a patient in these steep dose regions will be blurred compared to that calculated with the TPS or measured during one fraction. In specifying accuracy levels for IMRT treatments these additional uncertainties should be taken into account. However, also in 3DCRT steep dose gradients occur, which was not always fully realised in the past and generally assumed
to be taken into account in the PTV margin. With the introduction of IMRT people became more aware of the effect of movement on a calculated 3D dose distribution. It is therefore strongly recommended to deliver IMRT using a type of image guidance to assure the correct position of the target volume and/or organs at risk.

Similarly to the situation for 3DCRT, the intended dose distribution in the target volume should also be delivered for IMRT techniques with an uncertainty of 2 to 3%, one standard deviation, in the dose delivered to the dose specification point or volume, assuming a homogeneous dose distribution. In case more than one target volume is present, IMRT planning systems are able to optimise the dose distribution in these target volumes simultaneously. Such an approach is nowadays very often applied in case of simultaneous boost IMRT techniques. Also with the advent of image-guided radiotherapy it might be desirable to deliver non-homogeneous dose distributions to a target volume having, for instance, different hypoxic areas (e.g., Ling et al., 2000). For these situations accuracy requirements valid for homogeneous dose distributions have to be adapted by taking the volume of these relatively small regions with different dose levels into account.

7.2 TOLERANCE AND ACTION LEVELS OF TESTS FOR IMRT VERIFICATION

Tests for IMRT verification can be separated into those for verification of equipment for IMRT delivery, verification of IMRT treatment planning, and verification of patient-specific IMRT techniques, i.e., of the combined planning and delivery process of that particular patient treatment based on relative as well as absolute dosimetry. Proposed values for leaf position accuracy, leaf position reproducibility, gap reproducibility and leaf speed have been given by Palta et al. (2003) for both step-and-shoot and sliding window techniques. For instance, as tolerance limit for the leaf positioning a value of 1 mm was suggested for step-and-shoot treatments and 0.5 mm for the sliding window technique. The action level for each parameter is typically set at a value twice the tolerance limit. These authors also suggested limits for output stability for low MU delivery. These characteristics of accelerator and MLC performance are strongly dependent on the specific type of equipment (e.g., see Vieira et al., 2006). Some recent types of accelerators might easily fulfil these requirements, while older types might even with the best maintenance programme never be able to reach these goals due to limitations, for instance, in their mechanical construction. The proposed values for the tolerances for these parameters should therefore be considered as recommendations but in some cases not attainable in clinical practice. Obviously a decision has to be made during the commissioning process of the equipment what values should be chosen for the parameters applied during the QA process. Uncertainties in the characteristics of the accelerator and the MLC will also have a different influence on different treatment techniques. Each institution should therefore define its own tolerance limits and action levels for accelerator and MLC
performance, depending on the local situation, using for instance the data provided by Palta et al. (2003) as a useful starting point.

Verification of the specific IMRT issues of a treatment planning system has been discussed in Chapters 1 and 5 of this booklet. In Chapter 1 the experience of the QUASIMODO group has been elucidated, while in Chapter 5 the merits of independent dose calculations have been discussed in some detail. In these chapters two main approaches for IMRT verification are proposed. Either the combined planning and delivery process is verified experimentally, or an independent dose calculation is performed in combination with a thorough QC programme of accelerator and MLC performance. Both types of verification should be explored for each patient group (class solution), and a number of examples of both approaches are given in Chapter 6. Even if later on independent dose calculations are chosen as the preferred method of IMRT verification, at the start of the clinical implementation of a new IMRT technique it is strongly recommended to perform measurements of the 3D dose distribution delivered to a phantom and to compare these with planned dose distributions. It should be noted that even if the final decision to accept a plan is based on the verification of the 3D dose distribution of the total plan, it is also important to perform experimental verification of each delivered beam for a relatively large number of patients. In this way systematic errors can be traced that may otherwise be undetected due compensation of several effects (e.g., see the example given in Section 2.2).

The shift from an experimental patient-specific QA test to an independent dose calculation is a crucial point as it necessarily has important consequences in terms of time, organisation of the department and costs, as discussed for instance by Georg et al. (2007a,b). Such a decision will therefore also be influenced by non-scientific aspects such as the presence of sufficient personnel to perform the measurements and the reimbursement policy in a specific hospital/country. As a general rule, such a change should only be performed if sufficiently long experience in IMRT verification has been build up. This experience necessarily reflects various factors such as the number of IMRT patients treated, the local tools like the delivery and optimisation systems, the degree of complexity of the delivered fluences and the type of IMRT treatments. Good practice might be to gradually reduce the experimental patient-specific QA procedures in combination with an increase of QC of the machine/delivery systems and a concomitant implementation of an independent dose calculation method. In any case, it is very important to perform an extensive dosimetric verification procedure in situations when independent dose calculations do not give satisfactory results, for unusual treatment techniques and regularly for sample patients.

Different approaches exist for the comparison of sets of measured and calculated dose distributions, as discussed in Section 2.2 and in an earlier ESTRO booklet (Mijnheer et al., 2004). Each of these approaches needs, however, well-defined criteria for acceptance of a plan and procedures if these criteria are not met. Tolerance and action levels can be used, for instance,
in the same way as generally applied during QC measurements of linear accelerators. These quantities can be defined in the following way: whenever a parameter is found in the range below the tolerance level, the equipment is suitable for high quality radiation therapy. If, however, a parameter exceeds the action level, it is essential that appropriate actions be taken as soon as possible. Consequently, tolerance levels are appropriate limits for performance specification and for acceptance testing procedures, while action levels might be regarded as more relevant values for use in ongoing quality control activities. If a parameter has a value between the tolerance level and the action level, the responsible physicist will generally decide to continue with the treatment until a suitable moment for further investigation occurs. If such an investigation is not possible, then high quality treatments should no longer be performed with such equipment.

Tolerance and action levels should now be defined for the various tests to compare measured with calculated dose distributions. As discussed in Section 2.2 the most often applied dose evaluation techniques comprise a direct comparison of dose differences, a comparison of distance-to-agreement between measured and calculated dose distributions, and a combination of these two parameters: the gamma evaluation method. Traditional concepts to compare dose distributions in a more quantitative way subdivide the area of interest into regions with various dose gradients and apply different acceptance and tolerance criteria (e.g., Van Dyk et al., 1993). Such an approach is still useful for point dose measurements, for instance as applied during pre-treatment IMRT verification of the dose at the specification point applying ionisation chambers. From the findings described in Section 3.1 it can be concluded that differences of about 5% are generally significant for such a type of IMRT verification. Deviations larger than ± 5% should therefore firstly result in a review of the complete dosimetric procedure taking into account the various factors influencing the comparison result. If no explanation for the observed discrepancy can be given, the measurement may be repeated. A possible recommendation might then be that a tolerance limit of ± 3% and an action level ± 5% should be applied for these types of point dose verifications. These recommendations depend on the complexity of the type of IMRT treatment, and lower or higher tolerance and action levels may be more appropriate for simple or very complex IMRT techniques.

When the number of comparison points is large, simple methods of reporting deviations between dose measurements and calculations will collapse, and a method of compiling these deviations into a single number is required as a pass-fail criterion. Other methods have therefore been proposed, e.g., the use of the quantity “confidence limit” by Venselaar et al. (2001). The confidence limit is based on the average deviation between measurements and calculations for a number of data points in a comparable situation, and the standard deviation (SD) of the average of the differences. The confidence limit is then defined as the sum of the average deviation and 1.5 SD. The factor 1.5 was based on experience and a useful choice in clinical practice (Venselaar and Welleweerd, 2001). A multiplicative factor of 1.96 instead of 1.5 has later been proposed by Palta et al. (2003) for having 5% of the individual points
exceeding the tolerance level. For both the verification of individual beams, as well as for
the verification of patient-specific “hybrid” plans, Palta et al. proposed the set of values of
confidence limits and action levels for IMRT treatments as given in Table 7.4. These values
were arrived at using the results of an IMRT questionnaire that was mailed to 30 institutions
in the US that actively use IMRT. This questionnaire was designed to collect data on how
each institution views QA requirements and tolerance limits for IMRT planning and delivery.
An IMRT treatment plan should not be used clinically if the measured dose difference is more
than the value given as the action level, which serves therefore as a pass-fail criterion.

Table 7.4 Proposed values of the confidence limits and action levels for IMRT treatments (from Palta
et al., 2003).

<table>
<thead>
<tr>
<th>Region</th>
<th>Confidence Limit*</th>
<th>Action Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>High dose, low dose gradient</td>
<td>+/-3%</td>
<td>+/- 5%</td>
</tr>
<tr>
<td>High dose, high dose gradient</td>
<td>10% or 2mm DTA</td>
<td>15% or 3mm DTA</td>
</tr>
<tr>
<td>Low dose, low dose gradient</td>
<td>4%</td>
<td>7%</td>
</tr>
<tr>
<td>Dose fall off (d90-50%)</td>
<td>2mm DTA</td>
<td>3mm DTA</td>
</tr>
</tbody>
</table>

* The confidence limit is defined as the sum of the average deviation and 1.96 SD. The average
deviation used in the calculation of confidence limit for all regions is expressed as a percentage of the
prescribed dose according to the formula: 100% x (D_{calc} - D_{meas}/D_{prescribed}).

Application of the gamma evaluation method for selecting action levels is still in a develop-
ment stage. The following example illustrates its use in a clinical environment. De Martin
et al. (2006) applied the concept of the confidence limit when analysing gamma histograms
obtained during the verification of IMRT of head-and-neck treatments. Figure 7.1 shows the
flow diagram used by this group at Hospital San Raffaele in Milan for their decision protocol
applying various gamma evaluation criteria, including the confidence limit indicated by the
symbol γΔ, which is defined as: γΔ = γ_{mean} + 1.96 SD(γ).

Interestingly these authors found better results for treatments performed with their newly in-
stalled linac, probably due to a more accurate description of the beam penumbra in their TPS
as a result of using a diode instead of a relatively large ionisation chamber during dose profile
measurements. This observation indicates that a careful statistical analysis of patient-specific
verification data might reveal also systematic uncertainties valid for the whole patient group.
It also proves that decisions made during the commissioning process, in this case accepting
a certain beam profile in the TPS, have a direct impact on the outcome of the QA process.
It should be noted that the operator’s evaluation, in case the primary tolerance level is not
completely satisfied, includes an analysis of beam profiles as well as a more detailed look
at the position of points exceeding the tolerance level. Discussion of the operator with the
responsible physicist and/or radiation oncologist is a prerequisite for such an evaluation.
Figure 7.1 Example of a flow diagram showing a decision protocol based on a specific combination of gamma evaluation criteria (from De Martin et al., 2007).
Different values for the tolerance levels for dose difference and spatial accuracy in the gamma evaluation are applied clinically, adapted to the local situation in a specific centre. For example, Agazaryan et al. (2003) specified 3% and 3mm for the verification of composite IMRT plans, using an ionisation chamber for absolute dosimetry. Low and Dempsey (2003) typically use 5% and 2–3 mm as tolerance values of the 𝛾-distribution during their clinical evaluation. In a more recent study Childress et al. (2005) analysed about 850 films resulting from IMRT plan verification. Their results showed no dependence on energy, accelerator or treatment site, but varied for the different QA phantoms and treatment planning systems applied in that study. Their preferred gamma index tolerance criteria were 5% and 3 mm.

Analysis of gamma maps or gamma-area histograms provides a number of parameters that might be used for defining pass-fail criteria of a patient-specific IMRT plan. In addition to the confidence limit γΔ, as applied by De Martin et al. (2006), other parameters that are currently in use include average 𝛾 value, gamma angle, maximum 𝛾 value, where the maximum value should be based on a specific number, for instance 1% of the total number of gamma values, and the fraction of 𝛾 values above one, P>1. A number of groups, e.g., Childress et al. (2005) and Budgell et al. (2005) suggested P>1 as the main parameter to assess the agreement between measured and calculated dose distributions. Most likely a combination of these parameters should be used as action level based on clinically relevant criteria. For instance, in a recent paper van Zijtveld et al. (2006) used the mean value of γ, combined with the percentage of points with 𝛾 > 1, as a tool to make decisions based on clinically relevant criteria more straightforward. Stock et al. (2005) proposed a combination of three parameters, given in Table 7.5, in their acceptance protocol. Such a type of recommendation depends on the chosen set of gamma evaluation criteria, in this case a dose difference, relative to dose maximum, of 4% and a DTA of 3mm.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Average gamma</th>
<th>Maximum gamma</th>
<th>P&gt;1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptable</td>
<td>&lt; 0.5</td>
<td>&lt; 1.5</td>
<td>0 – 5%</td>
</tr>
<tr>
<td>Need further evaluation</td>
<td>0.5 – 0.6</td>
<td>1.5 – 2.0</td>
<td>5 – 10%</td>
</tr>
<tr>
<td>Not acceptable</td>
<td>&gt; 0.6</td>
<td>&gt; 2.0</td>
<td>&gt; 10%</td>
</tr>
</tbody>
</table>

However, many other combinations of gamma evaluation and acceptance criteria are possible as discussed by McDermott et al. (2007). For instance, one may choose for gamma evaluation criteria the combination of average gamma = 0.5, maximum gamma = 1.5 and P>1 = 2%, for a rather strict acceptance protocol. In a more relaxed protocol one may choose values of 0.8, 3 and 5% for these three parameters applying the same set of gamma evaluation criteria. The general procedure is to calculate first gamma maps for 2%/2 mm, 3%/3mm, 4%/4mm or
another combination of dose-difference and DTA criteria, and then determine the percentage of maps that passes. A “reasonable” choice of a specific combination of gamma evaluation and acceptance criteria should then be based on the accuracy of the applied measurement procedure, its workload, and the ability to detect problem areas in the intended dose distribution. Such an approach has been described in detail for the gamma evaluation procedure using film dosimetry as applied during the ESTRO QUASIMODO inter-centre IMRT verification project (Gillis et al., 2005). Combinations of gamma evaluation and acceptance criteria depend on many factors including the dosimetric equipment, calculation and measurement grid, and the data analysis software. The criteria might also be different for axial and coronal planes, because at planes far away from the isocentre different accuracies in the dose calculation might be present, for instance due to off-axis beam softening effects. It is therefore virtually impossible to provide general recommendations applicable for all situations, and the procedure outlined above should be followed to formulate a protocol applied for IMRT verification in a specific clinic. The statistical analysis of the results of a routine QA programme, possibly applied to a set of patients treated according to a class solution, can be very useful in defining appropriate tolerance/action levels taking into account the special aspects of IMRT relevant for a specific clinic. Furthermore it should be mentioned that the gamma evaluation method is not only applicable for the comparison of experimental data with dose calculations, but can also be used for comparing a clinical plan with an independent dose calculation or a Monte Carlo simulation. Some examples of decision protocols will be given in the next section.

It should be noted that it cannot a priori be assumed that the tolerance levels developed for 2D analysis are also valid for the evaluation of 3D dose distributions. Currently no recommendations for 3D dose evaluation are available and are therefore urgently needed. This is of special importance because new recommendations to be provided in a new ICRU document for reporting IMRT (ICRU, 2008), will no longer recommend specifying the dose at a single point, but as the dose to a specified percentage of the PTV. The report does not recommend any particular percent volume for a prescription, but in any case the median dose in the PTV should be reported. This quantity, D_{50\%}, the dose where the DVH curve of the PTV crosses the 50\% volume point, is in many situations close to the dose at the ICRU reference point, as well as to the mean dose in the PTV, thus making the transition easier. Future dose evaluation criteria should therefore include the correctness of D_{50\%}, in addition to spatial information about the actual dose delivery. Biological considerations, combined with the clinical experience from the 3DCRT era, may be required to develop tolerance and action levels for the evaluation of 3D dose distributions for an individual patient.

If a gamma evaluation exceeds a certain action level for a chosen combination of dose-difference and DTA criteria, then possible reasons for discrepancies such as variation in phantom positioning, linac performance and film analysis should first be investigated. If these experimental uncertainties are within accepted values, then it might be useful to repeat the expe-
riment to confirm the observed discrepancies. If the same areas of the gamma maps fail the tolerance criteria again, then these areas should be compared with the corresponding regions in the patient dose distribution, and the implications of such a failure should be discussed with the responsible physicist and radiation oncologist. For each patient a decision should then be made if a new plan has to be generated or if the differences are clinically acceptable. It should be noted that such patient-specific action levels depend on many decisive factors, including the position and size of the area that failed to pass the evaluation criteria, the dose level in the PTV or OAR, and the sensitivity of the plan for movement. Furthermore, gamma evaluation is currently mainly restricted to the dose delivered to the PTV, whereas the dose in an OAR is equally important. Extension of decision protocols including OARs is therefore urgently needed. The situation becomes even more complicated if dose distributions are evaluated in 3D, as discussed earlier.

As a final remark it should be noted that tolerance limits and action levels have proven to be very useful in everyday quality control of accelerators, but some parameters are not easily and quickly corrected or repaired and some may almost be impossible or very expensive to restore. On very rare occasions, it might therefore be justified to use the radiation equipment clinically, even if an action level has been exceeded. Such a delicate decision can only be taken after close consideration of the responsible physicist with the knowledge of clinicians and radiation technologists. For example, due to the non-perfect position of the leaves, there may be some small parts in the target volume having either an under- or over-dosage. If, due to mechanical tolerances, not all leaves can be adjusted to the recommended values, it still may be justified to perform some types of IMRT treatment if the effect of these imperfections on the total dose distribution is known. If this happens only once or a few times during 30 fractions, this may not be detrimental for the overall patient treatment. The decision to clinically use a treatment unit, in spite of the fact that an action level has been exceeded, has to be discussed thoroughly and documented for every treatment method. Under these special circumstances the action level can no longer be considered as restrictive, i.e. since the clinical relevance of a parameter can differ considerably from one treatment to another, it is impossible to implement an action level as a mandatory minimum demand. Moreover, the choice of not giving an IMRT treatment should also take into account a possible detrimental effect due to the sub-optimal alternative treatment (for instance by using a less conformal 3DCRT technique).

7.3 POSSIBLE PITFALLS AND POTENTIAL ERRORS TRACED BY IMRT VERIFICATION

As mentioned by several authors, and elucidated in the introduction and at several other places in this document, many potential sources of error exist when applying IMRT, some of these may be easily missed by people new to IMRT. A few pitfalls or cautionary notes are
sprinkled throughout this booklet, while several examples are given of errors detected by applying a dedicated QA programme or a specific QC device. In order to facilitate the drafting or modification of protocols for the verification of IMRT, a summary will be given in this section of examples of possible pitfalls and potential errors as experienced by the QUASIMODO group. It is not the purpose of this section to provide a complete overview of all possible errors that may occur in every centre applying IMRT; differences in equipment and in treatment techniques would require anyhow different approaches. It is merely the intention to give the reader of this report some ideas of what type of errors can be found and which possible actions could be taken to solve a specific problem.

**Type of error and suggested action**

- **Lacking algorithm in the TPS for tongue-and-groove effect.**  
  *Action*: Design and verify a new plan in which the tongue-and-groove effect plays a minor role. Discuss the issue with the TPS manufacturer.

- **Imperfect leaf calibration.**  
  *Action*: Recalibrate the leaf position and perform a new verification experiment.

- **Systematic deviations between TPS calculations and ionisation chamber measurements at the isocentre for plans with many small segments due to uncertainties in the output factor calculation.**  
  *Action*: Rescale the number of MUs. Discuss the issue with the TPS manufacturer.

- **Large regions with gamma values above one during repeated film measurements, while ionisation chamber measurements are correct.**  
  *Action*: Check if the film batch is not expired and if so repeat the measurement with a new batch.

- **Under-dosage in part of the PTV due to a wrongly positioned moveable bar in the couch top.**  
  *Action*: Take the actual bar position into account during the dose calculation.

- **Non-optimal commissioning of a TPS due to a less accurate description of the beam penumbra as a result of using a relatively large ionisation chamber during the dose profile measurements.**  
  *Action*: Use a smaller ionisation chamber or diode detector.

- **Over-dosage in regions between adjacent leaves because the higher radiation leakage in these regions compared to the centre of the leaves is not adequately taken into account in the TPS.**  
  *Action*: Perform measurements for representative clinical cases and quantify the effect. Discuss the issue with the TPS manufacturer.

- **Errors due to positioning a film in the region between two adjacent leaves.**  
  *Action*: Place the film under the leaves.

- **Errors at the start of the delivery of an IMRT treatment, *e.g.*, overshoot of the beam profile.**  
  *Action*: Adjust the start-up parameters of the accelerator.
• Larger errors during verification under actual delivery gantry angles than with measurements under zero degrees.
  
  Action: Position the measuring device on a holder attached to the gantry for field-by-field verification, or measure the total dose distribution under treatment conditions.

• An accumulation of relatively small errors, e.g., a 1.5% difference in linac calibration on the day of the measurement, and a non-optimal kernel in the dose calculation algorithm of the treatment planning system.
  
  Action: Repeat the commissioning of the TPS and perform regular checks of reference IMRT fields.

• Missing significant errors, e.g., resulting from MLC displacements, due to the limited resolution of the measuring device.
  
  Action: Move the device in different directions and repeat the measurement.

• Missing a segment by the dose calculation engine of the TPS.
  
  Action: Perform field-by-field measurements and check if all segments have been calculated and delivered. Discuss the issue with the TPS manufacturer.

• Missing errors at other parts of the PTV or in OARs by performing only one ionisation chamber measurement or an independent MU calculation at a point.
  
  Action: Perform also measurements in one or more planes for representative clinical cases.

• Wrong parameter in the TPS for the definition of leaf position.
  
  Action: Understand and verify the definition of leaf position in your TPS.

• Wrong plan sent to the accelerator.
  
  Action: Analyse the planning and data transfer process and try to reproduce and/or avoid the error.

### 7.4 DIFFERENT STRATEGIES FOR PATIENT-SPECIFIC IMRT VERIFICATION

In this section some examples will be given of different strategies for patient-specific verification of IMRT treatments based on the different approaches outlined in this report. All approaches assume a thorough QA programme of the irradiation equipment, including appropriate leaf position verification procedures, thus guaranteeing optimal machine performance. Furthermore it is assumed that at the introduction of a new IMRT technique information about the complete 3D dose distribution has been gathered, as discussed at the introduction of Chapter 3. The scenarios described in this section therefore mainly concern patient-specific verification of “class solutions”, site-specific treatment techniques where per patient only a limited number of degrees of freedom are varied, and limitations of the planning and delivery processes for a specific technique are known. Furthermore it should be realised that with growing experience in patient-specific QA, the confidence in the use of an IMRT technique will grow and a simpler QA protocol may be applied. Such a situation occurs particularly if
a large number of patients have been treated according to a class solution. It should be noted that until now most verification procedures are based on verification of the dose specified at the ICRU point in the PTV, whereas the ICRU in a forthcoming report is recommending specifying the dose in IMRT in a volume rather than at a point (ICRU, 2008). Incorporation of such a recommendation may influence future verification procedures.

The use of a flow diagram showing the various steps and decision criteria of such a protocol, as shown for instance in Figure 7.1, is recommended for all scenarios. Generally such a protocol is only valid in a particular clinic for a specific type of IMRT treatment and should therefore only be considered as starting point for application in another hospital. The protocol should also include a detailed description of the measurement and data analysis procedure. All protocols assume that sufficient knowledge is available about the possibilities and limitations of the treatment planning system. In other words, the commissioning and QA tests of the TPS as outlined in the various documents (e.g., IAEA, 2004, Mijnheer et al., 2004, NCS, 2006, IAEA, 2007) have been successfully performed, and uncertainties in the dose distribution calculated by the TPS are known. All procedures should be well-documented applying sheets having all details of the procedure; in a similar way as done for other QC measurements. Verification of IMRT delivery in a phantom does not provide information about the actual dose delivered to a patient, which requires an in vivo approach. Protocols for in vivo dosimetry are in principle similar to those for pre-treatment verification but will not be presented here because their use for IMRT verification is still limited. An example of an in vivo dosimetry protocol can be found in McDermott et al. (2007).

**Scenario-1:**

- An ionisation chamber measurement is performed to assess the absolute value of the total dose of all fields at one or several points.
- Values for the tolerance and action level, for instance ± 3% and ± 5%, respectively, should be defined for the verification of the total dose at these points.
- All separate fields are measured using film, at one depth in a special phantom, to verify relative dose distributions in one plane.
- A comparison between measured and calculated dose distributions is performed using 2D evaluation procedures, for instance using the γ-index method.
- Evaluation parameters of 2D dose distributions such as the confidence limit, maximum and average gamma value, and \( P_{>1} \), should be chosen.
- Various combinations of dose differences and DTAs, for instance those given in Tables 7.4 and 7.5, should be tested to obtain a reasonable compromise between workload and clinical acceptable action levels for individual fields.
- Examples in this booklet have been given in Sections 6.1.5 and 6.1.8.

**Scenario-2:**

- All separate fields are measured using a 2D detector array to verify relative dose distributions in one plane.
- A comparison between measured and calculated dose distributions is performed using the intrinsic software. The impact of the dimensions of the individual detectors should be minimised by combining several measurements with shifted position of the device or taken into account in another way, e.g., by using convolution methods.
- Evaluation parameters of multiple point data such as the confidence limit, maximum or average gamma value, or $P_{>1}$, should be chosen.
- Various acceptability criteria, for instance those given in Tables 7.4 and 7.5, should be tested to obtain a reasonable compromise between workload and clinical acceptable action levels for individual fields.
- If the system is unable to perform absolute dose verification, additional measurements with a calibrated ionisation chamber should be performed to assess the absolute value of the total dose of all fields at one or several points.
- Values for the tolerance and action level, for instance ± 3% and ± 5%, respectively, should be defined for the verification of the total dose at these points.
- Examples in this booklet have been given in Sections 6.1.2, 6.1.7 and 6.1.9.

Scenario-3:
- The (relative) total dose distribution of all fields is measured with film or a 2D array at one or several depths in a special phantom.
- A comparison between measured and calculated dose distributions is performed using 2D evaluation procedures, for instance using the $\gamma$-index method.
- Evaluation parameters of 2D dose distributions such as the confidence limit, maximum and average gamma value, and $P_{>1}$, should be chosen.
- Various combinations of dose differences and DTAs, for instance those given in Tables 7.4 and 7.5, should be tested to obtain a reasonable compromise between workload and clinical acceptable action levels for the total dose distribution.
- An ionisation chamber measurement is performed to assess the absolute value of the total dose of all fields at one or several points.
- Values for the tolerance and action level, for instance ± 3% and ± 5%, respectively, should be defined for the verification of the total dose at these points.
- Examples in this booklet have been given in Sections 3.2.1 and 7.1.

Scenario-4:
- The absolute value of the total dose of all fields is measured at one or several points in a special phantom using calibrated detectors.
- Values for the tolerance and action level, for instance ± 3% and ± 5%, respectively, should be defined for the verification of the total dose at these points.
- Examples in this booklet have been given in Sections 6.1.1 and 6.1.7.
**Scenario-5:**

- *The absolute value of the total dose of all fields is recalculated* using an independent dose calculation at one or several points using the patient CT data or CT data of a special phantom.

- Values for the tolerance and action level, for instance ± 3% and ± 5%, respectively, should be defined for the verification of the total dose at these points.

- A separate check is needed to verify the download of the leaf sequence file.

- *Examples in this booklet have been given in Sections 6.1.2, 6.1.3, 6.1.4 and 6.1.6.*
8. REFERENCES


APPENDIX A

LIST OF WEBSITES OF COMPANIES SELLING TOOLS FOR IMRT VERIFICATION

Capintec: www.capintec.com
CIVCO: www.civco.com
Computerized Imaging Reference Systems, CIRS: www.cirsinc.com
Eastman Kodak: www.kodak.com
Elekta: www.elekta.com
ESTRO-EQUAL: www.equal-dose.org
Euromechanics: www.euromechanics.de
Gammex rmi: www.gammex.com
International Specialty Products: www.ispcorp.com
Modus Medical Devices: www.modusmed.com ; www.quasarphantoms.com
Oncolog: www.oncolog.com
PTW-Freiburg: www.ptw.de
Radimage: www.radimage.com
ScandiDos: www.scandidos.se
Scanditronix-Wellhofer: www.scanditronix-wellhofer.com
Siemens: www.medical.siemens.com
Standard Imaging: www.standardimaging.com
Sun Nuclear: www.sunnuclear.com
Theraview: www.acceletronics.com ; www.cablon.n
Varian: www.varian.com