Evaluation of detector array technology for the verification of advanced intensity-modulated radiotherapy

by

Mohammad Hussien

A thesis submitted to the Department of Physics, University of Surrey, for the degree of Doctor of Philosophy

August 2015

©Mohammad Hussien 2015
Statement of originality

This thesis and the work to which it refers are the results of my own efforts. Any ideas, data, images or text resulting from the work of others (whether published or unpublished) are fully identified as such within the work and attributed to their originator in the text, bibliography or in footnotes. This thesis has not been submitted in whole or in part for any other academic degree or professional qualification. I agree that the University has the right to submit my work to the plagiarism detection service TurnitinUK for originality checks. Whether or not drafts have been so-assessed, the University reserves the right to require an electronic version of the final document (as submitted) for assessment as above.

© Mohammad Hussien 2015
# TABLE OF CONTENTS

**ABSTRACT** .......................................................................................................................... VII

**ACKNOWLEDGMENTS** ........................................................................................................ IX

**LIST OF FIGURES** ............................................................................................................... X

**LIST OF TABLES** .................................................................................................................. XV

**ABBREVIATIONS** .................................................................................................................. XVI

## 1 INTRODUCTION .................................................................................................................. 1

1.1 **BRIEF OVERVIEW OF RADIOTHERAPY** ........................................................................ 1

1.2 **OVERVIEW OF IMRT TECHNOLOGY** ........................................................................... 1

1.3 **ROTATIONAL IMRT** ....................................................................................................... 3

1.3.1 **Tomotherapy** ............................................................................................................ 4

1.3.2 **Intensity modulated arc radiotherapy (IMAT)** ......................................................... 5

1.3.3 **Volumetric modulated arc radiotherapy (VMAT)** .................................................... 5

1.4 **QUALITY ASSURANCE AND VERIFICATION FOR ADVANCED RADIOTHERAPY** ........... 6

1.5 **SCOPE OF THIS RESEARCH** ......................................................................................... 9

## 2 EVALUATION OF DOSE DISTRIBUTIONS: THE GAMMA INDEX ANALYSIS TECHNIQUE .......... 12

2.1.1 **Dose difference** ........................................................................................................ 12

2.1.2 **Distance-to-agreement** ............................................................................................. 13

2.1.3 **The composite index** ................................................................................................ 13

2.1.4 **The gamma index** .................................................................................................... 14

2.2 **LITERATURE REVIEW** .................................................................................................. 17

2.2.1 **Prevalence of the use of the gamma index** ............................................................... 17

2.2.2 **The computing challenge of the gamma index** ....................................................... 19

## 3 DETECTOR ARRAY TECHNOLOGY ....................................................................................... 22

3.1 **WHAT IS THE IDEAL DETECTOR ARRAY?** .................................................................... 22

3.2 **BRIEF HISTORY OF DETECTOR ARRAYS** ..................................................................... 23

3.2.1 **Review of detector array characterisation studies for IMRT and VMAT QA** .............. 26

3.3 **DESCRIPTION OF DETECTOR ARRAYS USED IN THIS THESIS** ................................ 28

3.3.1 **The PTW OCTAVIUS® Series** ............................................................................... 28

3.3.2 **Scandios Delta4®** .................................................................................................... 33

3.3.3 **SunNuclear ArcCHECK®** ....................................................................................... 34

3.4 **THE NEW TREND TOWARDS MEASUREMENT GUIDED 3D DOSE RECONSTRUCTION USING DETECTOR ARRAYS** ................................................................. 35

3.5 **SUMMARY** .................................................................................................................... 38

## 4 DEVELOPMENT OF METHODOLOGY TO CHARACTERISE DETECTOR ARRAY TECHNOLOGY .......... 40

4.1 **INTRODUCTION** ............................................................................................................ 40

4.2 **MATERIALS AND METHODS** ....................................................................................... 40
4.2.1 EBT2 Gafchromic Film .................................................................................................................. 40
4.2.2 Varian Electronic Portal Imaging Device (EPID) ........................................................................... 41
4.2.3 Comparisons of using the OCTAVIUS scan with 2D-ARRAY 729 in situ vs homogeneous scan .... 41
4.2.4 Multiple acquisition modes in the 2D-Array ................................................................................. 41
4.2.5 Deliberate plan modification tests ................................................................................................. 43
4.2.6 Gantry angle 0 degree test fields .................................................................................................... 44
4.2.7 Test clinical plans with deliberate modifications ............................................................................ 45
4.2.8 Effect of normalisation point .......................................................................................................... 46
4.2.9 Dosimetric and radiobiological evaluation of clinical plan modifications ................................. 47
4.2.10 Data and statistical analysis .......................................................................................................... 48
4.3 RESULTS ........................................................................................................................................ 49
4.3.1 Gantry angle 0 degree test fields ............................................................................................... 49
4.3.2 Dosimetric and radiobiological impact of the deliberate clinical plan modifications .................. 50
4.3.3 Composite verification of clinical plans ......................................................................................... 51
4.3.4 Comparison between expected gamma index passing rates in Verisoft and OmniPro ............... 52
4.3.5 Comparison between evaluations using CT scan with 2D-ARRAY 729 in situ vs homogeneous insert. 52
4.3.6 Single vs multiple 2D-ARRAY 729 acquisition modes ................................................................. 55
4.3.7 2D-ARRAY 729 vs Gafchromic film ............................................................................................. 56
4.4 DISCUSSION ................................................................................................................................ 57
4.5 CONCLUSIONS ............................................................................................................................. 59
5 EVALUATION OF THE IMPACT OF THE GAMMA INDEX CALCULATION APPROACH: A BESPOKE MATLAB SOFTWARE ............................................................................................. 60
5.1 INTRODUCTION ............................................................................................................................. 60
5.2 OPTIMISING THE SOFTWARE ......................................................................................................... 62
5.2.1 Impact of Limiting the Search Distance on Calculation Time ..................................................... 62
5.2.2 Data Interpolation techniques ..................................................................................................... 64
5.3 COMPARISON AGAINST TWO COMMERCIAL SOFTWARE .......................................................... 67
6 A COMPARISON OF THE GAMMA INDEX ANALYSIS IN VARIOUS COMMERCIAL IMRT/VMAT QA SYSTEMS ..................................................................................................................... 71
6.1 INTRODUCTION ............................................................................................................................. 71
6.2 DETECTOR ARRAY SYSTEMS INTER-COMPAED ........................................................................... 72
6.3 VIRTUAL PLAN MEASUREMENTS TO CALCULATE PREDICTED Γ IN COMMERCIAL SOFTWARE .................................................................................................................. 72
6.4 EXPERIMENTAL PLAN MEASUREMENTS ..................................................................................... 73
6.5 DATA AND STATISTICAL ANALYSIS ............................................................................................. 74
6.6 RESULTS ........................................................................................................................................ 74
6.6.1 Predicted γ based on virtual measurements .................................................................................. 74
6.6.2 Calculated γ based on experimental linac measurements .............................................................. 78
Contents

6.7 DISCUSSION .......................................................................................................................... 82

7 THE IMPACT OF THE GAMMA INDEX CALCULATION IN 2D AND 3D ON DETECTOR ARRAY RESOLUTION 

7.1 INTRODUCTION ..................................................................................................................... 85

7.2 METHOD .................................................................................................................................. 86

7.2.1 Equipment ............................................................................................................................. 86

7.3 DELIBERATE PLAN MODIFICATION TESTS .......................................................................... 86

7.4 RESULTS .................................................................................................................................. 87

7.5 DISCUSSION ............................................................................................................................. 90

8 DEVELOPMENT OF A METHODOLOGY FOR USING A DETECTOR ARRAY IN A NATIONAL ROTATIONAL 

IMRT DOSIMETRY AUDIT ............................................................................................................. 92

8.1 BACKGROUND AND RATIONALE ......................................................................................... 92

8.2 METHODS AND MATERIALS ................................................................................................. 93

8.2.1 The RTTQA 3DTPS Test for RapidArc, VMAT and Tomotherapy ........................................ 94

8.2.2 Choice of detector array and phantom combination ............................................................ 95

8.2.3 Semiflex ionization chambers ............................................................................................... 97

8.2.4 Gafchromic EBT2 film ......................................................................................................... 97

8.3 ALANINE PELLETS .................................................................................................................. 97

8.3.1 Miscellaneous equipment ..................................................................................................... 98

8.4 SEMIFLEX IONIZATION CHAMBER OUTPUT CALIBRATION ............................................... 98

8.5 2D-ARRAY CROSS-CALIBRATION ....................................................................................... 98

8.6 VERIFICATION PLAN CREATION ......................................................................................... 99

8.6.1 Measurement planes .............................................................................................................. 100

8.6.2 Individual dose point measurement .................................................................................... 102

8.7 ANALYSIS OF DATA .............................................................................................................. 103

8.7.1 Comparison of the 2D-Array against semiflex ion chamber and alanine ....................... 103

8.7.2 Comparison between 2D-Array and Gafchromic film planes ............................................ 103

8.7.3 Reporting of data to host centres ....................................................................................... 104

8.8 TIMING STUDY ..................................................................................................................... 104

8.9 RESULTS .................................................................................................................................. 104

8.9.1 Comparison of individual 2D-Array chambers against semiflex ion chamber and alanine .... 105

8.9.2 Comparison between 2D-Array and Gafchromic film ......................................................... 107

8.9.3 Timing study ....................................................................................................................... 109

8.10 DISCUSSION ........................................................................................................................ 110

9 A METHODOLOGY FOR DERIVING OPTIMAL GAMMA INDEX ACCEPTANCE CRITERIA FOR DETECTOR 

ARRAY MEASUREMENTS ............................................................................................................ 113

9.1 INTRODUCTION ..................................................................................................................... 113

9.2 METHODOLOGY .................................................................................................................... 114
Contents

9.2.1 Deliberate plan delivery deviation simulations ................................................................. 114
9.2.2 Retrospective analysis of measured data from the UK national audit of VMAT and Tomotherapy 115
9.2.3 Correlation of gamma index metrics against treatment plan DVH metrics .......................... 116
9.2.4 Correlation based on measured data .................................................................................. 116
9.2.5 Methods for Estimating Gamma Index Acceptance Criteria ............................................. 117
9.2.6 Comparison of different passing criteria .......................................................................... 119
9.2.7 Evaluation of different lower thresholds for the gamma index ........................................... 120
9.3 RESULTS ................................................................................................................................ 120
9.3.1 Correlation of gamma index metrics against DVH metrics .............................................. 120
9.3.2 Acceptance criteria based on ROC analysis ................................................................. 126
9.3.3 Acceptance Criteria Based on Confidence Limits ............................................................. 128
9.3.4 Correlation of different gamma index passing criteria .................................................... 128
9.3.5 The effect of varying the lower dose threshold for γ analysis ........................................... 129
9.4 DISCUSSION .......................................................................................................................... 130
10 DISCUSSION AND CONCLUSIONS ....................................................................................... 134
  10.1 DISCUSSION ...................................................................................................................... 134
  10.2 CONCLUSIONS ................................................................................................................... 136
11 FUTURE WORK .......................................................................................................................... 137
  11.1 INTRODUCTION ................................................................................................................ 137
  11.2 INVESTIGATION OF THE CHAPTER 9 METHODOLOGY IN VARIOUS COMMERCIAL DETECTOR ARRAY SYSTEMS ................. 137
  11.3 FURTHER EVALUATION OF DETECTOR ARRAYS FOR 3D DOSE GUIDED RECONSTRUCTION IN PATIENT ANATOMY .......... 137
  11.4 CORRELATION OF GAMMA INDEX AGAINST MACHINE SPECIFIC METRICS .................... 138
  11.5 CORRELATION OF GAMMA INDEX AGAINST TREATMENT PLAN SPECIFIC METRICS ...... 138
  11.6 CAN DETECTOR ARRAY DATA GIVE MEANINGFUL INFORMATION ABOUT THE ACCURACY OF LOW DOSE MODELLING? 139
  11.7 WHAT THRESHOLD VALUE SHOULD BE USED FOR ROC ANALYSIS TO DERIVE GAMMA INDEX ACCEPTANCE CRITERIA? 139
  11.8 EMERGING APPROACHES IN ADVANCED IMRT VERIFICATION .................................... 139
  11.9 INVESTIGATION OF OTHER EVALUATION METRICS ...................................................... 140
  11.10 EVALUATION OF THE OCTAVIUS1500 DETECTOR ARRAY FOR 3D DOSE RECONSTRUCTION ........................ 140
  11.11 CAN THE BESPOKE MATLAB SOFTWARE BE USED TO DETERMINE OPTIMUM DETECTOR ARRAY CONFIGURATION? 141
REFERENCES .................................................................................................................................... 142
APPENDIX A BASIC DETECTOR ARRAY COMMISSIONING TESTS ................................................... 158
  A.1 INTRODUCTION .................................................................................................................... 158
    A.1.1 Effective point of measurement ....................................................................................... 158
    A.1.2 Dosimetric linearity ........................................................................................................ 158
    A.1.3 Field size comparisons .................................................................................................. 159
    A.1.4 Comparisons of using the OCTAVIUS scan with 2D-ARRAY 729 in situ vs homogeneous scan ... 160
  A.2 RESULTS AND DISCUSSION .................................................................................................. 161
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.2.1</td>
<td>Basic commissioning</td>
<td>161</td>
</tr>
<tr>
<td>A.2.2</td>
<td>Effective point of measurement</td>
<td>161</td>
</tr>
<tr>
<td>A.2.3</td>
<td>Linearity of the 2D-ARRAY 729 with dose, dose-rate and dose-per-pulse</td>
<td>161</td>
</tr>
<tr>
<td>A.2.4</td>
<td>Directional response in the OCTAVIUS scan with 2D-Array in situ vs homogeneous scan</td>
<td>163</td>
</tr>
<tr>
<td>A.2.5</td>
<td>Output vs field size response of the 2D-Array</td>
<td>165</td>
</tr>
<tr>
<td>A.2.6</td>
<td>Comparison of the 2D-ARRAY 729 with diode profiles taken in the RFA Water tank</td>
<td>165</td>
</tr>
<tr>
<td>A.2.7</td>
<td>Comparison of the 2D-ARRAY 729 with Linear Diode Array (LDA) profiles taken in the RFA plotting tank</td>
<td>167</td>
</tr>
<tr>
<td>B.1</td>
<td>PUBLISHED RESEARCH PAPERS (ABSTRACTS)</td>
<td>169</td>
</tr>
<tr>
<td>B.1.1</td>
<td>Peer reviewed publications</td>
<td>171</td>
</tr>
<tr>
<td>B.1.2</td>
<td>Published conference abstracts</td>
<td>172</td>
</tr>
<tr>
<td>APPENDIX C</td>
<td>GAMMA INDEX SOFTWARE MATLAB CODE</td>
<td>173</td>
</tr>
</tbody>
</table>
ABSTRACT

Purpose
Quality assurance (QA) for intensity modulated radiotherapy (IMRT) has evolved substantially. In recent years, various ionization chamber or diode detector arrays have become commercially available, allowing pre-treatment absolute dose verification with near real-time results. This has led to a wide uptake of this technology to replace point dose and film dosimetry and to facilitate QA streamlining. However, arrays are limited by their spatial resolution giving rise to concerns about their response to clinically relevant deviations. The common factor in all commercial array systems is the reliance on the gamma index ($\gamma$) method to provide the quantitative evaluation of the measured dose distribution against the Treatment Planning System (TPS) calculated dose distribution. The mathematical definition of the gamma index presents computational challenges that can cause a variation in the calculation in different systems. The purpose of this thesis was to evaluate the suitability of detector array systems, combined with their implementation of the gamma index, in the verification and dosimetry audit of advanced IMRT.

Method
The response of various commercial detector array systems (Delta4®, ArcCHECK®, and the PTW 2D-Array seven29™ and OCTAVIUS II™ phantom combination, Gafchromic® EBT2 and composite EPID measurements) to simulated deliberate changes in clinical IMRT and VMAT plans was evaluated. The variability of the gamma index calculation in the different systems was also evaluated by comparing against a bespoke Matlab-based gamma index analysis software. A novel methodology for using a commercial detector array in a dosimetry audit of rotational radiotherapy was then developed. Comparison was made between measurements using the detector array and those performed using ionization chambers, alanine and radiochromic film. The methodology was developed as part of the development of a national audit of rotational radiotherapy. Ten cancer centres were asked to create a rotational radiotherapy treatment plan for a three-dimensional treatment-planning-system (3DTPS) test and audited. Phantom measurements using a commercial 2D ionization chamber (IC) array were compared with measurements using 0.125cm³ ion chamber, Gafchromic film and alanine pellets in the same plane. Relative and absolute gamma index ($\gamma$) comparisons were made for Gafchromic film and 2D-Array planes respectively. A methodology for prospectively deriving
Abstract

appropriate gamma index acceptance criteria for detector array systems, via simulation of deliberate changes and receiver operator characteristic (ROC) analysis, has been developed.

Results

In the event of clinically relevant delivery introduced changes, the detector array systems evaluated are able to detect some of these changes if suitable gamma index passing criteria, such as 2%/2mm, are used. Different computational approaches can produce variability in the calculation of the gamma index between different software implementations. For the same passing criteria, different devices and software combinations exhibit varying levels of agreement with the Matlab predicted gamma index analysis. This work has found that it is suitable to use a detector array in a dosimetry audit of rotational radiotherapy in place of standard systems of dosimetry such as ion chambers, alanine and film. Comparisons between individual detectors within the 2D-Array against the corresponding ion chamber and alanine measurement showed a statistically significant concordance correlation coefficient ($\rho$>0.998, p<0.001) with mean difference of -1.1%±1.1% and -0.8%±1.1%, respectively, in a high dose PTV. In the $\gamma$ comparison between the 2D-Array and film it was found that the 2D-Array was more likely to fail in planes where there was a dose discrepancy due to the absolute analysis performed. A follow-up analysis of the library of measured data during the audit found that additional metrics such as the mean gamma index or dose differences over regions of interest can be gleaned from the measured dose distributions.

Conclusions

It is important to understand the response and limitations of the gamma index analysis combined with the equipment and software in use. For the same pass-rate criteria, different devices and software combinations exhibit varying levels of agreement with the predicted $\gamma$ analysis. It has been found that using a commercial detector array for a dosimetry audit of rotational radiotherapy is suitable in place of standard systems of dosimetry. A methodology for being able to prospectively ascertain appropriate gamma index acceptance criteria for the detector array system in use, via simulation of deliberate changes and ROC analysis, has been developed. It has been shown that setting appropriate tolerances can be achieved and should be performed as the methodology takes into account the configuration of the commercial system as well as the software implementation of the gamma index.
ACKNOWLEDGMENTS

I would like to sincerely thank my supervisors Professor Andrew Nisbet and Dr Catharine Clark for their guidance and mentoring throughout. I am indebted to their wealth of experience and am grateful to their encouragement and allowing me the opportunities to explore my research ideas.

I am grateful to the Medical Physics Department of Royal Surrey County Hospital for funding my PhD fees, particularly Prof Nisbet for facilitating this and making it possible. Thanks to my Radiotherapy Physics colleagues, especially my line managers Tom Jordan, Liz Adams, and Steph Gerard-Martin for allowing me protected time from my routine clinical duties in order to be able to carry out my research.

Thanks to PTW Frieburg for loaning me the OCTAVIUS 4D phantom and the SRS1000 detector array. I would also like to thank Imaging Equipment Limited (IEL) and SunNuclear for loaning the ArcCHECK with SNC Patient software. I would like to thank Dr Sabeena Beveridge for assistance with some of the measurements for Chapter 6 and Yat Tsang and Mount Vernon Hospital for facilitating access to the flatbed scanner to process the Gafchromic film for the work in Chapters 4 and 6.

Thanks to Yat Tsang, from NPL: Russell Thomas, Julia Snaith, David Maughan, Clare Gouldstone, and from Christie Hospital: Steve Bolton, who collaborated on the audit methodology developmental visits. I would like to thank the staff at the audited centres for their hard work and cooperation. The audit methodology study was kindly sponsored by Varian Medical Systems Ltd, Elekta Systems Ltd and Oncology Systems Ltd (OSL, on behalf of Tomotherapy).

I would like to thank Dr Pejman Rowshanfarzad and Professor Martin Ebert, from Western Australia University, for the fruitful collaboration with the work in Chapter 6.

Thanks to the Institute of Physics and Engineering in Medicine (IPEM) for the award of a Travel Bursary to attend the ESTRO 33 Conference in Vienna 2014 to present some of my work.

Finally, I would particularly like to thank my wife, Safia, for all her love, encouragement and patience throughout. I would like to thank my family, parents and siblings, for all their love and support for all these years before and during the PhD, I wouldn’t be where I am now without this.
LIST OF FIGURES

FIGURE 1.1  COMPARISON OF CONVENTIONAL (3 FIELD), LEFT, AND AN IMRT (5 FIELD) PLAN, RIGHT ........................................... 2
FIGURE 1.2  SCHEMATIC OF IMRT DELIVERY FOR PROSTATE CANCER .................................................................................. 3
FIGURE 1.3  SCHEMATIC SHOWING DIFFERENCE BETWEEN STATIC IMRT (LEFT) AND ROTATIONAL IMRT (RIGHT). ................. 4
FIGURE 1.4  CORONAL IMAGE SHOWING THREE REGIONS (SHOWN AS BLUE RECTANGLES) FOR MEASURING IONISATION CHAMBER DOSE POINTS IN A CLINICAL PROSTATE & NODES IMRT TREATMENT PLAN. IN THIS EXAMPLE THE RECTANGLE DRAWN REPRESENT THE COLLECTION VOLUME OF A FARMER TYPE 0.6cm$^3$ IONIZATION CHAMBER AND ARE SAMPLING THE (1) HIGH DOSE PTV, (2) THE ELECTIVE NODAL PTV, AND (3) A LOW DOSE SPARING REGION ................................................................................. 8
FIGURE 1.5  SCHEMATIC REPRESENTATION OF TYPICAL SETUP FOR FILM IRRADIATION IN CUBIC SOLID WATER PHANTOM. THE FILM IS REPRESENTED AS YELLOW SHEETS SANDWICHED BETWEEN INDIVIDUAL LAYERS OF SOLID WATER ........................................................................................................ 8
FIGURE 1.6  SCHEMATIC REPRESENTATION OF FILM RESOLUTION (LEFT) VS DETECTOR ARRAY RESOLUTION (RIGHT). IN THE LATTER CASE, INDIVIDUAL DETECTORS ARE SHOWN AS GREY SQUARES. ......................................................................................... 9
FIGURE 2.1  DIAGRAMMATIC REPRESENTATION OF THE CONCEPT OF DOSE DIFFERENCE AND DTA. IN THIS EXAMPLE, THE REFERENCE DISTRIBUTION IS SHOWN AS THE BLUE LINE AND THE EVALUATED AS THE RED LINE ........................................................................................ 13
FIGURE 2.3  THE NUMBER OF CITATIONS PER YEAR (BLUE SOLID LINE) AND THE CUMULATIVE CITATIONS (BLACK DASHED LINE) FOR THE ORIGINAL GAMMA INDEX PAPER BY LOW ET AL [45]. .................................................................................................................. 18
FIGURE 3.1  SCHEMATIC SHOWING DIFFERENT TYPES OF DETECTOR ARRAY CONFIGURATIONS. PLANAR (LEFT), HELICAL ARRANGEMENT (MIDDLE) AND CROSS-PLANE (RIGHT) .............................................................................. 26
FIGURE 3.2  THE PTW OCTAVIUS II PHANTOM WITH 2D-ARRAY 729 DETECTOR ARRAY IN SITU. .................................................. 29
FIGURE 3.3  CT SCAN OF THE OCTAVIUS PHANTOM WITH (LEFT) 2D-ARRAY 729 IN SITU, AND (RIGHT) HOMOGENEOUS INSERT. ......................................................................................................................... 30
FIGURE 3.4  THE PTW OCTAVIUS DETECTOR 1000$^{60}$.................................................................................................................. 30
FIGURE 3.5  PROFILES ACQUIRED WITH THE BEAMNR/C DOSXY2NR/C MONTE CARLO SIMULATION CODE (MC), GAFCHROMIC FILM (GF), IONIZATION CHAMBER (IC), GLASS BEADS (GB) AND OPTICAL FIBRE (OF) FOR THE 10 X 10 CM, 4 X 4 CM, 3 X 3 CM, 2 X 2 CM AND 1 X 1 CM FIELD SIZES AT 5 CM WATER DEPTH AND NORMALIZED TO THE DOSE AT THE CENTRAL AXIS FOR A 10 X 10 CM FIELD SIZE DEFINED AT THE SURFACE. FROM JAFARI ET AL [95], WITH PERMISSION ......................................................................................................................... 31
FIGURE 3.6  THE OCTAVIUS4D PHANTOM SETUP ON A VARIAN CLINAC IX AT THE ROYAL SURREY COUNTY HOSPITAL . .......... 32
FIGURE 3.7  SCHEMATIC REPRESENTATION OF THE 3D DOSE RECONSTRUCTION ALGORITHM USED FOR THE OCTAVIUS 4D. THE RED LINE INDICATES THE MEASURED DOSE PLANE. THE BLUE LINES INDICATE THE RECONSTRUCTED DOSE PLANES AT INTERVALS ALONG THE RAY LINE FOR THE GANTRY ANGLE ......................................................................................................................... 33
FIGURE 3.8  THE SCANDIDOS DELTA4 DETECTOR ARRAY WHICH USES P-SI DIODES SETUP IN A BI-PLANE CONFIGURATION. ....... 34
List of figures

Figure 3.9 The SunNuclear ArcCHECK detector array which uses diodes set up in a spiral configuration. ....................................... 35

Figure 4.1 Schematic of the overlap between the 5mm Varian Millenium MLCs and ion chambers within the 2D-array for a collimator rotation of 0° (as typically used for static gantry IMRT) and 30° (as typically used for RapidArc). The dashed line indicates the central axis. .................................................................................................................. 43

Figure 4.2 Single gantry test fields for sensitivity (left) and resolution (right) assessment. In the sensitivity test the regions vary in width from left to right between 3mm and 15mm and the dose difference relative to the background varies from top to bottom by 1% to 10%. In the resolution test the values in the lower half represent difference in % dose between the regions and the background (lime green) area. In both tests, each row represents a single MLC leaf. ........................................................................................................................................ 45

Figure 4.3 Schematic diagram of cumulative histogram analysis for evaluating the gamma index response of different systems. On the Y-axis is a plot of the percentage of all measured distributions and on the X-axis is plotted the percentage of measured points with γ<1. For a range of measured plans with no delivery errors one would expect all measured distributions to achieve 100% of measured points to pass with γ<1 for a given passing criteria. If a range of plans with known delivery ‘errors’ are measured, one would expect an ideal system to demonstrate a range of failures and produce a trend such as that illustrated by the solid blue line. As such the response of different systems can be tested by comparing their results against a baseline trend. Measured trends that tend to the left of the baseline have greater response and those to the right have a lower response. This type of graph allows for a comparison between different systems and can be used to compare different passing criteria. ........................................................................................................................................ 48

Figure 4.4 (Left to right) Gafchromic film of highly modulated test, 2D-ARRAY 729 measurement (single acquisition), 2D-ARRAY 729 measurement (fully merged), EPID measurement. ............................................................... 49

Figure 4.5 Gamma index passing rates for the sensitivity test field for different measurement permutations.
Points have been linked to provide a visual guide. ........................................................................................................................................ 50

Figure 4.6 Expected vs measured local dose difference due to the MLC positional modifications. ......................... 51

Figure 4.7 (a) Cumulative histogram of gamma index passing rates at 3%/3mm predicted doses calculated using 2D-ARRAY 729 in situ and homogeneous scan. (b) Cumulative histogram of multiple acquisition analysis calculated at 3%/3mm (solid lines) and 2%/2mm (dashed lines). (c) Cumulative histogram of 2D-ARRAY 729 versus Gafchromic film gamma index analysis calculated at 3%/3mm and 2%/2mm. ........................................ 53

Figure 4.8 Gamma maps using 3%/3mm criteria showing the effect of multiple acquisition modes for prostate IMRT plan with 5mm MLC positional modification. (a) Predicted gamma index distribution; 99.5%, (b) Single acquisition; passing rate 96.8%, (c) Two merged acquisitions with array shifted lateral for second acquisition; 97.0%, (d) Two merged acquisitions with array shifted longitudinal for second acquisition; 96.2%, (e) Four merged acquisitions to give effective 5mm resolution; 96.8%, (f) Acquisition with 5mm shift in the longitudinal direction 96.1%. ........................................................................................................................................ 56

Figure 4.9 Comparison between gamma index distribution at a passing criteria of 3%/3mm for the head & neck IMRT plan with 5mm MLC positional modification (a) predicted, (b) 2D-ARRAY 729, (c) Gafchromic film and prostate & nodes plan with randomly distributed fluences modifications for (d) predicted, (e) 2D-ARRAY 729, (f) film. 57
List of figures

Figure 5.1 Screenshot of the gamma index calculation software implemented in MATLAB ............................................. 61
Figure 5.2 Gamma index map for the deliberate 5° collimator rotation change distribution against the normal
distribution, showing failed points expected in the 2.5 mm pixel spacing distribution ................................................. 63
Figure 5.3 Comparison of the maximum gamma index as a function of varying the search distance for global
gamma passing criteria 3%/3 mm, 3%/2 mm, 3%/1 mm, and 2%/2 mm. The average calculation time taken is
plotted against the right axis ........................................................................................................................................ 64
Figure 5.4 Comparison of different MATLAB interpolation algorithms; (a) spline, (b) linear, (c) cubic and (d)
nearest neighbour ....................................................................................................................................................... 65
Figure 5.5 Comparison of the maximum gamma index as a function of varying the interpolation factor for
global gamma passing criteria 3%/3 mm, 3%/2 mm, 3%/1 mm, and 2%/2 mm. The calculation time taken is
plotted against the right axis for each passing criteria .................................................................................................. 67
Figure 5.6 Gamma index map for the 10 mm resolution reference distribution from (a) OmniPro, (b) Verisoft, (c)
MATLAB plotted using the imagesc function, and (d) MATLAB plotted using the contourf function ..................... 68
Figure 5.7 Comparison of % points passing with $f<1$ between (top graph) MATLAB, Verisoft and OmniPro
analysis for reference distributions with 2.5 mm, 5 mm and 10 mm pixel spacing, and (bottom graph) repeated
using MATLAB with no interpolation setting. Analysis used global 3%/2 mm with no lower dose threshold ... 69
Figure 5.8 Comparison of mean gamma index between (top graph) MATLAB, Verisoft and OmniPro analysis for
reference distributions with 2.5 mm, 5 mm and 10 mm pixel spacing, and (bottom graph) repeated using MATLAB
with no interpolation. Analysis used global 3%/2 mm with no lower dose threshold ..................................................... 70
Figure 6.1 Cumulative histogram of gamma index analysis using passing criteria of 3%/3 mm for (a) comparison
of different software against the independent predicted calculation and (b) comparison of the measurements
by the QA systems against the independent prediction. The software in the figure legend in (a) is listed in the
same order as the associated hardware in (b) ................................................................................................................. 76
Figure 6.2 Scatter plot of gamma index analysis for each plan using passing criteria of 2%/2 mm for (a)
comparison of different software against the independent predicted calculation and (b) comparison of the
experimental measurements by the QA systems against the independent prediction. The dashed line gives a 1:1
trend that would be expected for perfect agreement. The software in the figure legend in (a) is listed in the
same order as the associated hardware in (b) ................................................................................................................. 77
Figure 6.3 Concordance correlation coefficients for the different systems for 3%/3 mm. For the Delta4, the
coefficient is also given for 4%/3 mm and 4%/4 mm ........................................................................................................ 78
Figure 6.4 Cumulative histogram of gamma index analysis for the Delta4 using passing criteria of 3%/3 mm,
4%/3 mm and 4%/4 mm against the predicted 3%/3 mm results .................................................................................. 79
Figure 6.5 Gamma index distributions for the head & neck IMRT plan with 5 mm MLC deliberate change. The
passing criteria shown is 3%/3 mm. Distributions are shown for the independent gamma index calculation, and
the measured plans using the 2D-Array, Gafchromic film, EPID, ArcCHECK and Delta4 ....... 80
Figure 6.6 Predicted vs measured gamma index points with $f>1$ for the ArcCHECK ............................................. 81
Figure 6.7 Percentage of detectors/pixels passing with a gamma index < 1, for 2%/2 mm, for the prostate IMRT
plan with 2 mm MLC positional modification, head & neck (H&N) IMRT plan with 2 mm MLC positional
List of figures

- Modification, prostate RapidArc (RA) with 2mm MLC positional modification, H&N RA with 2mm MLC positional modification and the prostate and pelvic nodes (PPN) RA plan with a collimator (col.) rotation error of 1 degree. ................................................................. 82
- Figure 7.1 The parameter specification window for 3D dose reconstruction in Verisoft V6.0. ......................... 87
- Figure 7.2 Global gamma index passing rates for 3%/2mm for the different equipment and simulated analysis. 89
- Figure 7.3 2D dose profile for the 3D reconstructed measured dose from the 729 and 1000\textsuperscript{SRS} arrays compared against the TPS dose profile. ........................................................................................................ 89
- Figure 7.4 2D coronal planes. Top row: TPS, SRS1000 in OCTAVIUS 4D, OCTAVIUS 729 in OCTAVIUS 4D, and 2D-ARRAY 729 in OCTAVIUS II dose distribution for the plan with 5mm MLC positional modification. Bottom row: 3%/3mm \( \Gamma \) distribution for the plan with 5mm MLC positional modification in Matlab predicted, OCTAVIUS-1000\textsuperscript{SRS}, OCTAVIUS 729, and coronal 2D-ARRAY. ................................................................. 90
- Figure 8.1 Transverse slice image of the 3DTPS test (left). The dashed lines indicate the planes measured. The black dots indicate dose points, note that two points were taken in PTV2 (one in the coronal and one in the sagittal plane). The image on the right shows a 3D reconstruction of the test with all PTVs and OAR labelled. 95
- Figure 8.2 The OCTAVIUS II system with (a) 2D-ARRAY 729, (b) Gafchromic EBT2 film insert, (c) PTW semisflex 0.125cc ion chamber insert, and (d) alanine insert. Note that for the 2D-ARRAY 729, the base of the OCTAVIUS phantom has a compensation gap to correct for the inherent under-response of the array; this is indicated by the black smile on the outside. For the other dosimetry systems, a solid base is used (indicated by the red smile). ........................................................................................................ 96
- Figure 8.3 Dose planes measured for the 3DTPS test .......................................................................................... 101
- Figure 8.4 Measuring plane dose distribution. 1\textsuperscript{st} coronal plane (a), 2\textsuperscript{nd} coronal plane (b), sagittal plane (c). 102
- Figure 8.5 Point dose measurements in (a) coronal 1, (b) coronal 2, and (c) sagittal plane. Alanine measurements were carried out in (a) and (c). ........................................................................................................ 103
- Figure 8.6 Scatter plots and Bland-Altman agreement plots for (a & b) individual 2D-ARRAY 729 detector dose against semisflex ion chamber measurements, (c & d) individual 2D-ARRAY 729 detector dose against alanine measurements, and (e & f) semisflex ion chamber measurements against alanine measurements. On each scatter plot, perfect agreement is indicated by the dashed line. On each Bland-Altman plot, upper and lower 95% limits of agreement are shown using dotted lines and mean difference using bold line. ......................................................... 106
- Figure 8.7 Relative frequency histograms showing (a) percentage difference between measured dose and predicted dose for the 2D-Array, semisflex and alanine, and (b) 2D-ARRAY 729 versus Gafchromic film gamma index analysis calculated at 3%/3mm. ......................................................... 108
- Figure 8.8 Comparison of the gamma index distribution between the 2D-ARRAY 729 and Gafchromic film in a case where the absolute point dose difference was <1% and a case where the dose difference was 4%. In the former case it can be seen there is good agreement between the 2D-ARRAY 729 and film. In the latter case, the difference between the absolute and relative gamma index is evident. ......................................................... 109
List of figures

**Figure 8.9** Median timescale for audit visits, assuming a 9:00 AM start time. .................................................. 110

**Figure 9.1** Example gamma index distribution for a 3 mm left-right alignment deviation. Calculated using global gamma 3%/2 mm. ........................................................................................................... 115

**Figure 9.2** 3%/2 mm global 3D volume (a) gamma index pass-rate, (b) mean gamma, (c) median gamma, (d) max gamma, (e) near-max gamma against the PTV2 mean absolute dose deviation. ............................................... 121

**Figure 9.3** 3%/2 mm global 3D volume (a) pass-rate, (b) mean gamma, (c) median gamma, (d) max gamma, (e) near-max gamma against OAR maximum dose deviation. ................................................................. 122

**Figure 9.4** Simulated 3D volumetric gamma index vs 2D coronal and sagittal planes for (left) mean gamma index and (right) passing rate for 3%/2 mm, 20% threshold. ............................................................. 124

**Figure 9.5** 3%/2 mm, 20% threshold global gamma index passing rate (points passing with \( \Gamma < 1 \)) given as a percentage (plotted on the left y-axis), median, mean and near-max gamma index (plotted on the right y-axis) plotted against (a) the mean absolute dose difference in the 2D ROI corresponding to PTV2 mean and (b) the absolute maximum dose difference in the 2D ROI corresponding to OAR. The trend line for the \( \Gamma_{\text{max}} \) is not displayed. 125

**Figure 9.6** Simulation study ROC analysis of the sensitivity and specificity of gamma metrics to predict a 3% deviation in (top) PTV2 mean dose deviation and (bottom) OAR maximum dose deviation. Gamma analysis performed using 3%/2 mm with 20% threshold. ........................................................................................................... 127

**Figure 9.7** Simulated (a) and measured (b) mean gamma index values for passing criteria 3%/3 mm and 2%/2 mm against 3%/2 mm. The dashed line indicates a 1:1 agreement. ................................................................. 129

**Figure 9.8** Average (a) \( \Gamma_{\text{mean}} \) (b) gamma index passing rate for the measured data as a function of lower dose threshold. The error bars represent standard error of the mean. The line linking the individual points is for visualization only. ........................................................................................................... 130
LIST OF TABLES

TABLE 2.1 List of top 15 citing journals for the gamma index.  
18

TABLE 3.1 Summary of characteristics of the available commercial detector arrays (as of March 2015).  
25

TABLE 4.1 Summary of mean and minimum gamma index passing criteria for all various measurement permutations. A lower number indicates greater response to deviation detection. The concordance correlation coefficient, \( p_c \), is also given assessing agreement with single 2D-array acquisition.  
54

TABLE 5.1 Comparison of different interpolation techniques using global 1%/1mm, 0% threshold.  
66

TABLE 6.1 Summary of the mean and minimum measured gamma index passing criteria for each system. The concordance correlation coefficient, \( p_c \), is also given assessing agreement with independent gamma index. The software are listed in the same order as the associated measurement system.  
75

TABLE 7.1 Summary of mean and minimum gamma index passing criteria for each system for the range of changes introduced. The concordance correlation coefficient, \( p_c \), is also given assessing agreement with the predicted 3D volume gamma index pass-rate.  
88

TABLE 8.1 Summary of equipment and delivery techniques in the ten centres visited.  
94

TABLE 8.2 3DTPS treatment planning objectives.  
95

TABLE 8.3 Range of detectors/pixels passing with \( \Gamma < 1 \) for the 2D-Array and Gafchromic film planes for passing criteria 4%/4mm, 3%/3mm, and 2%/2mm. The percentage of 2D-Array and Gafchromic film planes achieving \( \geq 95\% \) of detectors or pixels with a \( \Gamma < 1 \), \( p_{95\%} \), is also given.  
107

TABLE 9.1 Summary of mean, coefficient of variance (\( c_v \)) and Pearson correlation coefficient (\( p \)) for each \( \Gamma \) parameter for 3%/3mm, 3%/2mm, 2%/2mm using a 20% threshold. All \( p \) have \( p<0.01 \).  
123

TABLE 9.2 Summary of mean, coefficient of variance (\( c_v \)) and Pearson correlation coefficient (\( p \)) for each \( \Gamma \) parameter for 3%/3mm, 3%/2mm, 2%/2mm using a 20% threshold. All \( p \) have \( p<0.01 \).  
124

TABLE 9.3 Acceptance criteria for the gamma index metrics to predict a \( \pm 3\% \) dose deviation in the PTV2 mean dose and OAR maximum dose based on the simulation and measured data. In all cases, sensitivity was 100%.  
128

TABLE 9.4 Acceptance criteria based on confidence interval calculations for the measured audit data for Global 3%/2mm, 20% threshold.  
128

TABLE 9.5 Acceptance criterion based on ROC analysis for the measured audit data for Global 3%/2mm, 20% threshold for PTV dose deviation.  
131
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D</td>
<td>Two dimensional</td>
</tr>
<tr>
<td>3D</td>
<td>Three dimensional</td>
</tr>
<tr>
<td>3DCRT</td>
<td>Three dimensional conformal radiotherapy</td>
</tr>
<tr>
<td>AAA</td>
<td>Analytical Anisotropic Algorithm</td>
</tr>
<tr>
<td>AAPM</td>
<td>American Association of Physicists in Medicine</td>
</tr>
<tr>
<td>CT</td>
<td>Computed Tomography</td>
</tr>
<tr>
<td>DICOM</td>
<td>Digital Imaging and Communications in Medicine</td>
</tr>
<tr>
<td>DTA</td>
<td>Distance-to-Agreement</td>
</tr>
<tr>
<td>DVH</td>
<td>Dose Volume Histogram</td>
</tr>
<tr>
<td>EPID</td>
<td>Electronic Portal Imaging Device</td>
</tr>
<tr>
<td>EQUAL</td>
<td>European Quality Assurance Network</td>
</tr>
<tr>
<td>ESTRO</td>
<td>European Society for Radiotherapy and Oncology</td>
</tr>
<tr>
<td>FFF</td>
<td>Flattening Filter Free</td>
</tr>
<tr>
<td>IAEA</td>
<td>International Atomic Energy Agency</td>
</tr>
<tr>
<td>IC</td>
<td>Ion chamber</td>
</tr>
<tr>
<td>ICRU</td>
<td>International Commission on Radiation Units &amp; Measurements</td>
</tr>
<tr>
<td>IMAT</td>
<td>Intensity Modulated Arc Radiotherapy</td>
</tr>
<tr>
<td>IMRT</td>
<td>Intensity Modulated Radiotherapy</td>
</tr>
<tr>
<td>IPEM</td>
<td>Institute of Physics and Engineering in Medicine</td>
</tr>
<tr>
<td>Linac</td>
<td>Linear Accelerator</td>
</tr>
<tr>
<td>MLC</td>
<td>Multi-Leaf Collimator</td>
</tr>
<tr>
<td>MV</td>
<td>Mega-Voltage</td>
</tr>
<tr>
<td>NCRI</td>
<td>National Cancer Research Institute</td>
</tr>
<tr>
<td>NHS</td>
<td>National Health Service</td>
</tr>
</tbody>
</table>
Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPL</td>
<td>National Physical Laboratory</td>
</tr>
<tr>
<td>NTCP</td>
<td>Normal tissue complication probability</td>
</tr>
<tr>
<td>OAR</td>
<td>Organ at Risk</td>
</tr>
<tr>
<td>PDD</td>
<td>Percentage depth dose</td>
</tr>
<tr>
<td>PDP</td>
<td>Planned Dose Perturbation</td>
</tr>
<tr>
<td>PMMA</td>
<td>Polymethyl methacrylate</td>
</tr>
<tr>
<td>PTV</td>
<td>Planning Target Volume</td>
</tr>
<tr>
<td>QA</td>
<td>Quality Assurance</td>
</tr>
<tr>
<td>QC</td>
<td>Quality Control</td>
</tr>
<tr>
<td>RSCH</td>
<td>Royal Surrey County Hospital</td>
</tr>
<tr>
<td>RTTQA</td>
<td>Radiotherapy Trials Quality Assurance</td>
</tr>
<tr>
<td>SABR</td>
<td>Stereotactic Ablative Body Radiotherapy</td>
</tr>
<tr>
<td>TCP</td>
<td>Tumour control probability</td>
</tr>
<tr>
<td>TERMA</td>
<td>Total energy released per unit mass</td>
</tr>
<tr>
<td>TPR</td>
<td>Tissue Phantom Ratio</td>
</tr>
<tr>
<td>TPS</td>
<td>Treatment Planning System</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>VMAT</td>
<td>Volumetric Modulated Arc Radiotherapy</td>
</tr>
<tr>
<td>VODCA</td>
<td>Visualization and Organization of Data for Cancer Analysis</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
</tbody>
</table>
1

INTRODUCTION

1.1 BRIEF OVERVIEW OF RADIOThERAPY

Radiotherapy is the use of ionising radiation to treat benign and malignant (cancerous) superficial and deep-seated tumours within the human body. Throughout history, cancer has been a major cause of morbidity, but is now the second biggest cause of death in the western world. It has been estimated that one in two people in the UK born after 1960 will develop cancer at some stage in their lives due to increased life expectancy [1]. More than 200 different types of cancer have been identified to date, all cancers however are basically similar; they all result from uncontrolled cell growth causing tumours. The four most common cancers, which make up over 50% of all cancer incidences, are breast, lung, bowel (including anus) and prostate cancer. Radiotherapy plays an important role in curative and palliative treatment of cancer, sometimes in conjunction with other therapeutic agents such as surgery, chemotherapy and hormonal therapy.

Until recently the standard technique for treating cancer using radiotherapy in the UK was Three Dimensional Conformal Radiotherapy (3DCRT). This technique uses 3-4 beams of uniform intensity shaped around the tumour. As technology advanced so had the ability to develop more accurate techniques of delivering radiotherapy. Advanced techniques such as Intensity Modulated Radiation Therapy (IMRT) and more recently Volumetric Modulated Arc Radiation Therapy (VMAT) have recently been introduced into clinical practice and are quickly becoming the ‘gold standard’. These techniques use multiple beams of varying intensity therefore better conforming the high doses to the tumour and resulting in better normal tissue avoidance.

1.2 OVERVIEW OF IMRT TECHNOLOGY

Intensity-Modulated Radiation Therapy can be typically described as ‘a radiation treatment technique with multiple (radiation) beams in which at least some of the beams are intensity-modulated and intentionally deliver a non-uniform intensity to the target. The desired dose distribution in the target is achieved after superimposing such beams from different directions. The
additional degrees of freedom are utilised to achieve a better target dose conformity and/or better sparing of critical structures’ [2]. A lower dose is therefore delivered to normal tissues around the tumour target. Figure 1.1 shows a comparison between a three-dimensional conformal radiotherapy (3DCRT) prostate treatment plan and an IMRT treatment plan. It can be clearly seen that the 3DCRT plan results in a high dose to the rectum whereas better sparing is achieved in the IMRT plan. Better conformity to the target can also be seen in the IMRT plan.

![Image of IMRT and 3DCRT plans](image)

**Figure 1.1** Comparison of conventional (3 field), *left*, and an IMRT (5 field) plan, *right*.

The early history of the concept of IMRT began in 1982 when Brahme et al [3] published their landmark paper on ‘Solution of an integral equation encountered in radiation therapy’. The paper looked at an idealised case with rotational symmetry. This was a doughnut-shaped target volume with a circular critical structure in the middle. The integral solution calculated the beam intensity profile that delivers the doughnut shaped distribution through a full rotation of the beam and therefore showing that a highly non-uniform intensity profile was needed to produce a uniform dose distribution in the doughnut-shaped target volume.

The practical formulation of the inverse solution was developed by Webb [4]. The inverse problem has no exact solution as there are no physical intensity modulated profiles that deliver the full prescription dose to the tumour and no dose to surrounding critical structure. It is desirable to therefore design an IMRT treatment plan such that it comes as close as possible to the ideal prescription. This has become the basis of the optimisation solutions used in modern IMRT planning.

Before the advent of multi-leaf collimators many researchers and investigators of IMRT in the late 1980’s/early 1990’s assumed that many more than 10 beams were required to simulate a rotational treatment. Bortfeld et al. [5] realised that less than 10 radiation fields were often enough to provide
clinically acceptable dose conformity and OAR sparing and therefore his work paved the way for motor-driven MLCs to be practically implemented in IMRT.

There are currently two distinct methods of delivering static gantry IMRT. This collaborative PhD was based at the Royal Surrey County Hospital NHS Foundation Trust (RSCH) where the sliding window technique is used; see Figure 1.2 for a schematic. In the sliding window technique the radiation beam is constantly on while the MLCs move across the field at a defined speed such that the intensity of the beam is modulated. The MLC leafs conventionally move from left to right. The sliding window IMRT is typically used by the Varian Linear Accelerators (Varian Medical Systems, Palo Alto, CA). At other centres, IMRT may be delivered using the step-and-shoot method, typically found on Elekta Linear Accelerators (Elekta Ltd, Crawley, UK).

![Figure 1.2 Schematic of IMRT delivery for prostate cancer.](image)

1.3 **ROTATIONAL IMRT**

The term rotational IMRT may be used to classify techniques that deliver a modulated dose with a constantly rotating gantry. The concept of delivering radiation treatment through an arc existed years before the introduction of 3DCRT and IMRT. Conformal modulated arc therapy was first suggested in 1965 [6]. In the last 30 years, significant developments such as computer-controlled
MLCs, IMRT, and the ever increasing power of desktop computer processing speed and memory have been achieved to bring the technology from research labs and into the clinic. These include Tomotherapy [7], Intensity Modulated Arc Radiotherapy (IMAT); [8], and Volumetric Modulated Arc Radiotherapy (VMAT) [9].

1.3.1 Tomotherapy

In Tomotherapy [7], the linear accelerator head is mounted in a CT scanner style gantry. Radiation is delivered as a rotating fan-beam. Intensity modulation is achieved by means of a binary collimator which either opens or closes. In the initial research system, a slice-by-slice approach was used; however this required precise indexing of the table from one slice to the next slice. Tomotherapy was subsequently developed as a helical delivery system. This included the development of a slipping rotating gantry to achieve more efficient dose delivery by simultaneous gantry rotation and couch translation. Tomotherapy’s triumvirate of a rotating fan beam, couch translation, and binary MLCs allows for highly complex distributions to be achieved. In addition, long volumes can be treated. For example it has been demonstrated that Total Bone Marrow Irradiation can be achieved successfully by conforming the high dose to the marrow whilst sparing surrounding normal tissue [10]. On a conventional Linac, a comparable plan would require multiple isocentres.
1. Introduction

1.3.2 Intensity modulated arc radiotherapy (IMAT)

IMAT was proposed as an alternative to Tomotherapy [8] using a conventional ‘C-arm’ linear accelerator. IMAT uses a rotating cone-beam with varying shapes and dose weightings to achieve modulation. The full arc is first approximated as a series of evenly spaced fixed fields, yielding multiple intensity patterns. The strategy in IMAT is to convert those patterns into multiple segments delivered with multiple overlapping arcs.

As the gantry moves from one angle to the next, MLC speed restrictions must be taken into account; i.e. an optimal field shape may need to be altered to allow for a smooth delivery. This may reduce plan quality in complex situations [11]. IMAT does not need to move the patient during treatment and also allows for non-coplanar delivery.

Although IMAT was first demonstrated in 1995 [8], the concept did not advance commercially and there was limited clinical uptake. There may be a few reasons for this. Firstly, there was no efficient optimisation algorithm available and the burden of optimising individual segments within the arcs would have been computationally expensive, and therefore time consuming. Secondly, the inability of linacs, at the time, to vary dose-rate dynamically meant that equally spaced angles had to be used and combined with the technical limitations of the MLC, plan quality for complex cases was degraded. Thirdly, it was necessary to deliver IMAT through multiple overlapping arcs, which may have been unattractive compared to static IMRT due to the increased delivery time.

However, it had been predicted that a single arc with a sufficient number of aperture shape variations could create an optimal treatment plan [8]. Various works have been performed that have attempted to use a single arc for IMAT and have been reviewed [11]. The major commercial breakthrough came after Otto [9] developed a single-arc IMAT algorithm by assuming that the dose-rate could be varied. This was referred to as volumetric modulated arc therapy (VMAT).

1.3.3 Volumetric modulated arc radiotherapy (VMAT)

In VMAT, the dose rate and gantry speed variation as well as MLC modulation is performed within the optimisation algorithm [9]. This uses progressive beam angle sampling to optimise a large number of apertures (~1 aperture / 2° gantry angle). Aperture shapes and weights are optimised initially for a number of coarsely spaced gantry angles and, as the solution converges, additional gantry angles are inserted. In the initial steps, aperture connectivity is ignored, giving the algorithm more freedom to find an optimal solution. As the angular spacing becomes finer, aperture shapes
are linearly interpolated from their angular neighbours. The result is a smooth transition from one aperture shape to the next meaning the gantry can be continuously rotating. This allows for delivery time within 2 minutes for a standard 2Gy/# treatment regime. This contribution and others [8,9,12–24] has paved the way for a rapid clinical and commercial uptake.

Varian Medical Systems adopted Otto’s VMAT solution and marketed it with the trademark name RapidArc™ in 2007 (Varian Medical Systems Ltd, Palo Alto, CL). The main change made by Varian was that the Linac control was updated to allow variable dose rate and gantry rotation. Shortly after, Elekta (Crawley, UK) marketed their solution with the trademark name VMAT™. Various third party manufactures have created a VMAT solution that can be used on either Varian or Elekta machines, for example Philips Medical Systems developed a VMAT algorithm known as SmartArc™ [25].

A search for peer-reviewed articles in Scopus using the keywords Volumetric Modulated Arc Therapy, RapidArc, SmartArc, returned 822 literature citations between 1st January 2008 and 31st December 2014. In 2008 there were just 7 papers but by the end of 2010 there were 122, and 228 by the end of 2011. This dramatic increase in the number of publications within a short space of time (particularly when one considers the timescales involved in the peer review process, which can take a few months for a paper to be published) demonstrates the remarkable speed of the clinical uptake of this technology.

For a clinical perspective on VMAT, the reader is referred to the comprehensive review that has been conducted by Teoh et al [26]. This gives a detailed account of the clinical use of VMAT in a range of different cancer sites. A systematic review of treatment planning studies has been published in the journal Radiotherapy & Oncology [27] which the reader is also referred to.

1.4 QUALITY ASSURANCE AND VERIFICATION FOR ADVANCED RADIOTHERAPY

The advancing rate of radiotherapy technology necessitates the need for extra quality assurance [28]. As the complexity of treatment technology continues to increase so do the potential uncertainties and inaccurate dose delivery can have clinical implications [29,30]. Additionally, radiotherapy dosimetry audits allow for the testing of procedures and the identification of errors [31–43].

The complexity of the 3D dose distributions in IMRT treatments requires careful quality assurance. IMRT distributions are characterised by numerous steep dose gradients in order to conform as
1. Introduction

tightly as possible to the target volume whilst minimising the dose to normal tissue. Conventional 3DCRT treatments are composed of relatively large uniform beams and therefore patient-specific quality assurance consisted of simple independent dose and monitor unit verification calculations which are supplemented by routine machine specific QA (which includes basic checks such as output constancy, energy, beam flatness & symmetry etc.). In IMRT, the complex MLC pattern means that an independent dose calculation alone is not sufficient as these methods, even at the time of writing, estimate a point dose. In IMRT, the dose at any point is delivered by a fraction of the total modulated field. The MLC pattern varies from patient to patient and the number of MUs is heavily linked to the complexity of this pattern. Therefore it is necessary to perform a patient-specific QA measurement to verify the fluence from the IMRT beams to ensure the suitability of the MLC pattern. Traditionally this has been done using ionisation chambers and film within cubic or semi- anthropomorphic phantoms. An example is given in the next paragraph, however the interested reader is referred to the review by Low et al [44] which gives a detailed account of these early IMRT QA methods.

Once an IMRT treatment plan for a patient is complete it is possible in radiotherapy treatment planning systems to create a verification plan. Essentially, a verification plan is a copy of the same geometry, dynamic MLC, and monitor units calculated on a CT scan of the physical phantom to be used for performing the verification measurement. This plan can be used to generate a predicted dose for comparison against the measurement. For ionisation chamber measurements, the predicted dose can be typically calculated by creating a contour on the CT dataset that simulates the collecting volume of the chamber (see Figure 1.4). The mean dose to this structure is then recorded. It was often necessary to move the isocentre position relative to the phantom in order to position the ion chamber in regions of the dose distribution that are of interest, e.g. in the high dose PTV region. It is also necessary to set the position so that the dose across the chamber is homogeneous and to avoid areas of high dose gradient as a minor error in the setup of the phantom can result in large dose measurement error. For comparison against film measurements (see schematic in Figure 1.5), a dose plane can be exported from the TPS at the same plane as the film within the phantom.
1. Introduction

Figure 1.4 Coronal image showing three regions (shown as blue rectangles) for measuring ionisation chamber dose points in a clinical prostate & nodes IMRT treatment plan. In this example the rectangle drawn represent the collection volume of a Farmer type 0.6cm$^3$ ionization chamber and are sampling the (1) high dose PTV, (2) the elective nodal PTV, and (3) a low dose sparing region.

Figure 1.5 Schematic representation of typical setup for film irradiation in cubic solid water phantom. The film is represented as yellow sheets sandwiched between individual layers of solid water.

In this example, the ion chamber provides an absolute point dose measurement and the film provided a relative 2D measurement of the IMRT fluence. These methods however are time consuming, particularly for film measurements which require heavy resources for calibration, processing, and analysis as well as the costs of single-use films requiring multiple batches to be purchased, which also require physical archiving. Radiographic or radiochromic film suffers from variation in the sensitivity from one batch to another and non-linear dose response resulting in measurement uncertainties that render them unsuitable for routine absolute dose checks. These resource costs have historically limited the number of patients that could be treated with advanced IMRT. Therefore a new measurement methodology was needed.

In recent years, various commercial 2D and 3D ionization chamber or diode detector arrays have become available. These electronic devices have allowed for verification of absolute dose in 2D or
1. Introduction

3D with near real-time results. This allows for analysis to be performed in the IMRT QA measurement session and therefore out of tolerance results can be investigated immediately. At the time of writing, conventional methods such as ionization chamber point dose measurements and film dosimetry are gradually being replaced by detector arrays.

International recommendations, such as from the European Society for Radiotherapy & Oncology (ESTRO), advise that 3D measurements for pre-treatment QA should be performed for every patient. The International Commission on Radiation Units and Measurement (ICRU) recommends that for low gradient regions (<20% per cm) of the dose distribution, the dose difference normalised to the prescribed point should be no more than ±3.5% and for high gradient regions (>20%/cm) the dose points should have a distance to agreement of ≤3.5mm. However the most common quantitative technique used is called the gamma index analysis which combines dose difference and distance to agreement into a single dimensionless metric [45]. A gamma index of less than 1 indicates that the measurement point lies within the dose difference and/or distance to agreement passing criteria. A common acceptance threshold is that ≥95% of measured points should pass with a gamma index of <1 for passing criteria 3% dose difference and 3mm distance to agreement [46]. Further information on the theory of this metric is given in Chapter 2.

1.5 Scope of this Research

Detector arrays are limited by their detector spacing, as shown in Figure 1.6, giving rise to concerns about their sensitivity to errors. Understanding the limitations of these devices is therefore critical.

Figure 1.6 Schematic representation of film resolution (left) vs detector array resolution (right). In the latter case, individual detectors are shown as grey squares.
1. Introduction

The aim of this PhD is to conduct a critical appraisal of this new measurement technology for the verification of IMRT and VMAT. The main research questions are:

1. How do detector arrays respond to known changes in treatment plan delivery?
2. How is the gamma index calculation affected by the limited detector spacing and the different available commercial configurations?
3. Are there software and/or hardware, or both, effects on the gamma index calculation?
4. Can a detector array be used in a dosimetry audit in place of standard methods such as ion chamber and film?
5. What optimal tolerances should be used?

The following gives an outline of the topics discussed in the different chapters of this thesis:

- The theoretical background behind the gamma index method of quantitative evaluation in IMRT QA, which is a significant theme throughout this thesis, is the focus of Chapter 2.
- Chapter 3 gives a review of detector array technology and the various options that are available, and those used in this thesis.
- A methodology developed for characterising and critically evaluating commercial detector arrays for IMRT/VMAT pre-treatment verification is developed in Chapter 4. This work has been published in the Journal of Applied Clinical Medical Physics. This chapter addresses research question number 1 by using clinical treatment plans with deliberately introduced changes.
- Chapter 5 introduces a bespoke gamma index software written in Matlab (MathWorks Inc.) by the author which has been developed to investigate the computing challenges of the gamma index calculation and to look at the impact of different software implementations on its calculation. Chapter 5 addresses research question number 2.
- Chapter 6 focuses on a comparison between three commercial detector array systems in terms of the combined hardware and software effect on the gamma index calculation. This work has been published as a research paper in Radiotherapy & Oncology. It has also been presented orally at the 2nd ESTRO Forum in Geneva 2013 and as an invited talk in the IPEM-RTSIG IMRT Verification Meeting in London 2012. Chapter 6 addresses research question number 3.

2 Hussein M et al. Radiotherapy & Oncology 2013;109:370-6. Copyright was required to be transferred to Elsevier for publication; however permission has been granted by Elsevier to reproduce work in this thesis.
1. Introduction

- The effect on 2D and 3D gamma index calculations of the spacing of detector arrays is investigated and discussed in Chapter 7. This work has been partially presented as poster presentation at the ESTRO 33 Conference in Vienna April 2014. Chapter 7 addresses research question numbers 2 and 3.

- Work has been carried out to develop a methodology for a National Rotational Radiotherapy Audit in the UK in collaboration with the National Physical Laboratory (NPL), National Cancer Research Institute Radiotherapy Trials Quality Assurance group (NCRI RTTQA), and Institute of Physics and Engineering in Medicine (IPEM). The details of this work are discussed in Chapter 8. This work has been published as an original research paper in Radiotherapy & Oncology and has been presented orally at the ESTRO 31 conference in Barcelona 2012, and the IPEM Biennial Radiotherapy Meeting in Oxford 2012. Chapter 8 addresses research question numbers 4.

- Chapter 9 focuses on methodologies for deriving optimal acceptance criteria for detector arrays for use in benchmarking studies. Chapter 9 addresses research question numbers 4.

- Chapter 10 gives the overall discussion and conclusions

- Chapter 11 discusses the prospects for future research directions respectively.

A full list of papers and presentations arising from this work are listed in Appendix B.

---

3 Hussein M et al. Radiotherapy & Oncology 2013;108:78-85. Copyright was required to be transferred to Elsevier for publication; however permission has been granted by Elsevier to reproduce work in this thesis.
This chapter reviews the different methods for comparing dose distributions that have been published in the literature. Initially, these methods were developed to be able to compare measured water phantom data for basic photon fields against 3D treatment planning system calculations as part of the commissioning process. The methods have then naturally been adopted in IMRT QC. This chapter focuses in detail on the most commonly used technique, the gamma index [45]. The following published terminology is used to distinguish between the two dose distributions that are being compared [45]:

- The **reference dose distribution** is generally taken as the ‘gold standard’, e.g. it could be the dose distribution that has been measured. In theory this could be a single point measurement, 1D (e.g. a line profile), 2D (e.g. film measurement), or 3D (e.g. gel dosimetry, Monte Carlo simulation).

- The **evaluated dose distribution** is what is being compared. In most cases this will be the predicted TPS dose distribution that is being checked for accuracy.

### 2.1.1 Dose difference

The most basic evaluation method is to take the dose difference between the two datasets and check if each point agrees within a specified tolerance value. This works well in regions where there is a low dose profile gradient. However, in regions of high dose gradient (as would be expected with IMRT) techniques, this evaluation method can report a large dose error due to a small misalignment which may not be clinically significant; see schematic representation of this in Figure 2.1. The dose difference is calculated using equation 2.1:

\[
\Delta D(r_R, r_E) = \frac{D_E(r_E) - D_R(r_R)}{D_R(r_R)}
\]  

(2.1)
Where \( D_E(r_E) \) is the dose at a point in the evaluated dose distribution, \( r_E \), and \( D_R(r_R) \) is reference point dose.

2.1.2 Distance-to-agreement

The distance-to-agreement (DTA) is another simple method that can be used. Ideally the spatial discrepancy between two distributions needs to be known and this tool was developed for such a purpose [47]. The DTA is the closest distance between the reference point and the point in the evaluated distribution that has the same dose value. In contrast to the dose difference test, the DTA test works better in high dose gradient regions where it can be interpreted as the spatial offset between the two distributions. It should be noted that the offset could be due to experimental setup error rather than inaccuracy of the evaluated dose distribution. DTA is very sensitive in low dose gradient regions. As shown in Figure 2.1, it is possible in a shallow dose region for the distance between two dose points of the same value to be large, and lead to a clinically insignificant error as the dose difference in that region may be small.

2.1.3 The composite index

As outlined above, the dose difference and DTA tests have advantages and disadvantages associated with them. The dose difference test is useful in low dose gradient regions but overly sensitive in high dose gradient, whereas the opposite is true for DTA. Therefore the ‘composite index’ was developed.
as a combination of the two tests, such that one can test both the dose and spatial agreement of two distributions, see Equation 2.2. The user sets pass/fail criteria for the dose difference and DTA.

\[
\text{Composite Index } (r_R) = \begin{cases} 0, & DTA < \delta r \cup \Delta D(r_R,r_E) < \delta D \\ 1, & DTA < \delta r \cap \Delta D(r_R,r_E) < \delta D \end{cases}
\] (2.2)

Where \(\delta D\) is the dose difference criterion and \(\delta r\) is the distance difference criterion.

### 2.1.4 The gamma index

The gamma index \((\gamma)\) evaluation [45] has become a standard technique used to evaluate measured distributions in commercial detector systems against the dose distribution predicted by commercial treatment planning systems. It combines \(\delta D\) and \(\delta r\) to calculate a dimensionless metric for each point in the evaluated distribution. A \(\gamma\) of < 1 indicates that points lie within the \(\delta D/\delta r\) passing criteria. In a QA scenario, the total percentage of points that have achieved \(\gamma<1\), for a given \(\delta D/\delta r\) criteria, is calculated and a pass/fail threshold is set [48].

![Diagrammatic representation of the gamma index method in 1D. Adapted from Low et al (1998).](image)

The y-axis is Dose, \(D\), and the x-axis is distance, \(r\). In this schematic, the reference point is illustrated by the cross at the origin of the graph. The blue line represents the evaluated dose distribution with individual points represented by blue dots.
2. The gamma index theory

In practice the gamma index is calculated based on finding the minimum Euclidean distance for each reference point, see Figure 2.2. For each reference point in the dose distribution, calculate against each point in the evaluated distribution:

1. the distance between reference to evaluated point: $\Delta r(r_R, r_E)$
2. the difference between the reference and evaluated dose; $\Delta D(r_R, r_E)$

Then for each point in the evaluated distribution, calculate the gamma index using the following equation:

$$\Gamma(r_R, r_E) = \sqrt{\frac{\Delta r^2(r_R, r_E)}{\delta r^2} + \frac{\Delta D^2(r_R, r_E)}{\delta D^2}}$$

Where $\delta r$ is the distance criterion and $\delta D$ is the dose difference criterion.

The gamma index is then taken as the minimum value calculated over all evaluated points.

$$\gamma(r_R) = \min\{\Gamma(r_R, r_E)\} \forall \{r_E\}$$

An array of gamma index values for all of the points within the reference dose distribution can be constructed. It is then common to report the percentage of points passing with a $\gamma < 1$. This is called the gamma index passing rate. For nomenclature it is standard to report the passing criteria in the format $\delta D(\%) / \delta r(mm)$; for example, 1%/1mm. This standard nomenclature is used throughout this thesis. In order to eliminate dose in the out-of-field region where a large relative dose difference can be calculated and skew the gamma index result, it is typical to set a lower dose threshold below which the gamma index result is ignored. Therefore, it is common to limit the gamma index calculation to all points that are $\geq 10-20\%$ of the maximum dose value within the dose distribution [49].

2.1.4.1 Global & local gamma index calculations

Typically the gamma index calculations are categorised into two different types; local and global. The contrast between the two types is the way the dose difference is calculated. For a local gamma index, Equation 2.1 gives the definition for a local dose difference. For global gamma, Equation 2.1 has to be modified to become:
2. The gamma index theory

\[ \Delta D(r_R, r_E) = \frac{D_E(r_E) - D_R(r_R)}{D_{\text{norm}}} \]  \hspace{1cm} (2.5)

Where \( D_{\text{norm}} \) is a normalisation dose value which can be defined as any value, and is usually defined as the maximum dose within the reference dose distribution or a point selected in a high dose low gradient region. The two types of gamma index reporting have advantages and disadvantages. The local gamma index will tend to highlight failures in high dose gradient regions and in low dose regions, whereas the global gamma index will tend to mask these errors but show the errors within the higher dose regions within the dose distribution. The choice of the type of gamma index will depend on the need of the test. Most published works within the reference list report global gamma index. In most of this thesis, much of the focus is on global gamma index, and is discussed further in the next few chapters.

2.1.4.2 Advantages and disadvantages of the gamma index

It is common to report the results of a gamma index analysis as the number points that achieved gamma index <1; i.e. the gamma index passing rate. The main characteristic of this metric is that it can condense a verification measurement into a single unit; this is both the advantage and pitfall of this metric. If implemented carefully, it can be used to streamline QA by making it possible to choose decision thresholds for a passing rate, thereby reducing the analysis time. The disadvantage is that the passing rate does not provide any details of where failed points are. Another disadvantage is that the gamma index itself is inherently an absolute metric, i.e. it provides no information on whether a failed point is due to positive or negative dose or distance fluctuations. For example, it is possible for there to be a failed point in an OAR region where the measurement is lower than the predicted dose by more than the dose difference criterion. In this case, the failed point is clinically acceptable as the OAR is receiving a lower than expected dose and the aim of radiotherapy is to keep dose to an OAR region as low as possible. Conversely, it is possible to have a failed point in a PTV region where the measured dose is higher which would also be acceptable.

The most common passing criteria being used is 3%/3mm which was originally recommended in the work by Low et al [45]. The gamma index was originally designed to compare measured water tank beam data against a treatment planning system algorithm. The criteria of 3%/3mm were used due to the limitations of computing capability at the time (1998).
2. The gamma index theory

The choice of appropriate passing criteria are investigated and discussed further throughout this thesis. The potential limitations of the gamma index are also discussed further.

2.2 Literature review

2.2.1 Prevalence of the use of the gamma index

The prevalence of the use of the gamma index was investigated through a literature search. Unfortunately a search on Scopus and PubMed for “gamma & index & radiotherapy” returned some papers that were unrelated to radiotherapy measurement as the use of “gamma” is common in other scientific fields such as cancer biochemistry. However, all uses of the index in the literature should have cited the original publication by Low et al in Medical Physics in 1998 and therefore Scopus was used to investigate citations to this article and what time trend there has been since its initial publication. A search on PubMed was attempted for this article but gave no citations before 2006 and therefore was abandoned.

In the Elsevier Scopus abstract and citation database, (as of March 2015) it was found that the gamma index paper has been cited 865 times in the literature since it was published. Of these, there were 775 original research articles; the remainder were composed of 65 conference proceedings, 19 review papers and the remainder as book chapters or Editorials. The breakdown by scientific journal shows Medical Physics publishing ~33% of all research papers that made use of the gamma index. The top 15 citing journals which represent some of the major journals in the field are shown in Table 2.1. This shows that the use of the gamma index is not limited to specialist medical physics journals but also journals such as Radiotherapy & Oncology which are aimed at a wider community which also includes other members of the radiotherapy multidisciplinary team such as clinicians and radiographers.
2. The gamma index theory

Table 2.1 List of top 15 citing journals for the gamma index.

<table>
<thead>
<tr>
<th>Journal</th>
<th>Citing articles</th>
<th>% of total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical Physics</td>
<td>258</td>
<td>33.3</td>
</tr>
<tr>
<td>Physics in Medicine and Biology</td>
<td>140</td>
<td>18.1</td>
</tr>
<tr>
<td>Journal of Applied Clinical Medical Physics</td>
<td>74</td>
<td>9.5</td>
</tr>
<tr>
<td>Radiotherapy and Oncology</td>
<td>41</td>
<td>5.3</td>
</tr>
<tr>
<td>International Journal of Radiation Oncology Biology Physics</td>
<td>40</td>
<td>5.2</td>
</tr>
<tr>
<td>Medical Dosimetry</td>
<td>26</td>
<td>3.4</td>
</tr>
<tr>
<td>Journal of Physics Conference Series</td>
<td>26</td>
<td>3.4</td>
</tr>
<tr>
<td>Radiation Oncology</td>
<td>24</td>
<td>3.1</td>
</tr>
<tr>
<td>Australasian Physical and Engineering Sciences in Medicine</td>
<td>16</td>
<td>2.1</td>
</tr>
<tr>
<td>Zeitschrift Fur Medizinische Physik</td>
<td>14</td>
<td>1.8</td>
</tr>
<tr>
<td>Physica Medica</td>
<td>14</td>
<td>1.8</td>
</tr>
<tr>
<td>Journal of Medical Physics</td>
<td>12</td>
<td>1.5</td>
</tr>
<tr>
<td>British Journal of Radiology</td>
<td>11</td>
<td>1.4</td>
</tr>
<tr>
<td>Acta Oncologica</td>
<td>11</td>
<td>1.4</td>
</tr>
<tr>
<td>Strahlentherapie Und Onkologie</td>
<td>11</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Figure 2.3 The number of citations per year (blue solid line) and the cumulative citations (black dashed line) for the original gamma index paper by Low et al [45].
2. The gamma index theory

2.2.2 The computing challenge of the gamma index

The gamma index is a computationally expensive process due to the need to search all points in the evaluated distribution. This becomes more complex when comparing two 3D dose distributions. Ideally the gamma index would be calculated quickly to give a result within a reasonable time. The computer hardware used will have an impact on the speed of the \( \gamma \) calculation. Given previous limitations with computer technology, a number of studies in the literature have focussed on ways to mathematically decrease the calculation time of the gamma index. A common feature when reading through manufacturer manuals is that there is seldom detailed information on the way the gamma index calculation has been performed, with often the ‘go-to’ solution being to simply cite the original paper by Low et al. A true gamma index calculation can take from a few minutes up to potentially days for complex 3D dose distributions, which is clearly unacceptable from a clinical point of view [50]. This section gives a brief description on some of the methodologies and also discusses the future trend in this area.

2.2.2.1 Mathematical techniques to refine/speed up the gamma index calculation

A refinement of the gamma index was proposed by Depuydt et al [48]. In their work, the authors introduced a filter cascade of multiple levels that were designed to speed up the comparison. The focus of this algorithm is whether or not a point passes or fails the gamma index criteria rather than calculating an absolute gamma index value for each point in the reference distribution. The algorithm employs 3 levels in the calculation. In the first level, the minimum gamma index is searched within a limited distance to reduce the calculation time. As soon as an evaluated point is found where \( \Gamma(r_R, r_E) < 1 \) then the calculation is stopped and the algorithm moves on to the next reference point, starting at Level 1. If there are no evaluated points where \( \Gamma(r_R, r_E) < 1 \) then the algorithm moves on to Level 2. In the second level, the algorithm searches for at least 2 evaluated points in the vicinity of the reference point where the \( \Delta \Gamma(r_R, r_E) \) for the two points is of opposite signs. In this scenario it is assumed that the evaluated distribution must intersect the region defined by the passing criteria and therefore the reference point is classified as passed. Failing this, the algorithm moves to the third and last level. In most cases, reference points rejected in Level 2 because the evaluated points are truly outside the pseudo-space defined by the passing criteria. However, there may be rare occasions when the reference points should not have failed because of the discrete nature of the evaluated distribution. This is because it is possible for two discrete evaluated points at the outside edge of the passing region to produce a failed result, but interpolating between them means that there is an intersection with the passing region. Therefore
2. The gamma index theory

Level 3 is designed to take these possibilities into account. If the reference point fails level 3 then it is classified as having failed the test.

Chen et al [50] consider the possibility to speed up the search distance algorithm by using fast Euclidean distance transform and predict a speedup of the order of tens of thousands for 3D gamma index calculations. Bakai et al [51] published a revision of the original gamma index formalism which considered gradient-dependent local acceptance thresholds. The actual number calculated by the Bakai method is called the $\chi$ index and is defined using the following equation:

$$\chi = \frac{D_E(r_E) - D_R(r_R)}{\sqrt{\delta r^2 + \delta D^2} \nabla D_R(r_R)}$$ (2.6)

In this method, one begins by calculating the local gradient for each point in the reference distribution, $\nabla D_R(r_R)$, to build a gradient map or cube. The numerator part of equation 2.6 means that this method only works when the reference and the evaluated dose distributions have the same array sizes. This formalism means that a $\chi$ distribution can be calculated efficiently using simple matrix operations which are optimised for speed in numerical analysis coding software such as Matlab, or could be programmed in other languages using existing libraries and utilising multiple threads in the computer processing unit (CPU). The $\chi$ index retains the sign, unlike the gamma index. The value of $|\chi|$ is also approximately equivalent to the gamma index [51].

Wendling et al [52] developed a fast algorithm through speeding up the search routine by pre-sorting the distance from the reference point to evaluated points within a fixed search circle or sphere, for 2D and 3D respectively. The theory is that the search loop should be stopped when there is a low chance that evaluated data points will reach the minimal $\Gamma$. In their approach, the calculation starts at the reference point and increases outwards and terminating when the condition defined by equation 2.7 is met [52]:

$$\frac{\Delta r^2(r_E,x_E)}{\delta r^2} > \min(\Gamma(r_R,r_E))$$ (2.7)

The possible limitation of this technique is that there is potential for overestimation of the gamma index when dose differences are very large within the search region and then sharply drop off just outside this region. If an exhaustive search was performed in this scenario it would likely find the minimum gamma index in this region where the calculation has stopped.

Ju et al [53] propose a re-interpretation of the gamma index to avoid the need to interpolate the evaluated distribution to a finer grid, thus reducing the calculation time. For 1D, 2D, and 3D
distributions the evaluate distribution is divided into line segments, triangles and tetrahedral respectively. The closest distance between any reference point and these simplexes is calculated using matrix multiplication and inversion. The finding is that the method is as accurate as 16 times finer linear interpolation of the evaluated dose distribution with an order of ~20 times speedup in calculation time.

2.2.2.2 Future trends in computational techniques

As hardware continues to progress, the speed of the gamma index calculation will naturally continue to be quicker. In the last few years there have been innovative approaches to utilising the processing power provided by graphical processing units (GPU). Traditionally programs are calculated on the CPU, however the number of cores is limited with standard desktop PCs at the time of writing offering dual core processors and quad core at the higher price range. However, GPUs have been designed to handle complex graphics particularly in modern 3D video games where intensive tasks such as rendering and in-game physics simulations are required. The GPU is otherwise left unused when graphics intensive programs are not running. The parallel computing platform, called Compute Unified Device Architecture (CUDA), was developed by NVIDIA Corporation (Santa Clara, California, USA) which was first made available on compatible NVIDIA GPU cards in 2006 [54]. The CUDA platform can be implemented in standard programming languages such as C, C++ and Fortran, and software such as Matlab has implemented support for it. OpenCL™ (Open Computing Language, developed by Apple Inc., California) is a similar platform that has the advantage in not being vendor specific, although CUDA can work on AMD (Advanced Micro Devices Inc., California) GPU cards which represent the other major manufacturer. Driven by the video gaming industry, GPU cards are now routinely available with highly parallel multiple cores that are in the region of 1000+. This allows for significant potential for parallel computing in areas such as Monte Carlo calculations or in this case, fast gamma index calculations. For the gamma index, it would be possible to utilise the GPU to perform the otherwise computer intensive task of the minimum distance search by searching and calculating $\gamma$ for each reference in parallel.

Gu et al [55] studied accelerated gamma index calculations by combining the geometric technique proposed by Ju et al [53] and the pre-sorting technique described above by Wendling et al [50] and by implementing the calculation onto the GPU. They found a 45 – 70 times speedup of the calculation compared to the traditional implementation on the CPU. Peerson et al [56] found a speedup of $57 \pm 15$ for patient cases when using the GPU against the CPU.
3 DETECTOR ARRAY TECHNOLOGY

3.1 WHAT IS THE IDEAL DETECTOR ARRAY?

It is best to start by thinking about what would make an ideal detector array. There are many systems that have been researched and investigated in the literature that ultimately aim to come as close as possible to the ideal condition. The following is essentially a wish list of criteria for an ideal detector:

1. Have a high resolution which is comparable to or better than the resolution of the grid spacing used for the dose calculation in the TPS
2. Be able to measure a true 3D dose distribution; i.e. detectors arranged in a 3D lattice
3. Have no angular dependence
4. Have linear dose, energy and dose rate response
5. Be water equivalent
6. Be a robust system; i.e. suffer from no leakage during measurement sessions and have perfect short and long term reproducibility
7. Be easy to calibrate
8. Be able to perform real-time measurements so that a diagnosis of out-of-tolerance events can be made immediately

Of course all of the above cannot be easily achieved in a cost-effective manner. The only dosimeters currently available that have been able to measure a true 3D dose distribution are polymer gel dosimeters. The advantage of these is that they are tissue-equivalent and can be moulded into an anthropomorphic shape. After irradiations the dosimeter requires scanning using MRI, optical CT, or X-ray CT and then processing of the measured signal. It has been estimated that the entire process from fabrication to analysis can take up to 45 hours, rendering this unsuitable for routine
3. Detector array technology

measurement [57]. It has potential to be used as a benchmarking tool for a commissioning treatment plan. Currently this technology has been mainly confined to research institutions and the current processing and analysis timescales have meant that there is a limited market. However with research into optimisations and more cost-effective scanning techniques this may become more available in the future [57].

Manufacturers have attempted to fulfil as many of the above wish list by developing electronic detector arrays within the limits of technology and cost. Engineering challenges with building a 3D detector array include ensuring that the detectors respond equally from any direction in addition to ensuring that the circuitry required causes negligible perturbation in the dose. The following section outlines the available commercial electronic detector arrays and those that were investigated in this thesis.

3.2 Brief history of detector arrays

The first commercial detector array was the MapCHECK® (Sun Nuclear Corp., Melbourne, FL) which was a 2D array utilising 445 n-type diodes [52]. The diodes had a spacing of 7.07 mm in the central 10 cm x 10 cm area and 14.14 mm in the outer regions up to an area of 22 cm x 22 cm. This has since been superseded by the MapCHECK® 2 which increased the number of detectors to 1527 and maintained a uniform spacing of 7.07 mm to have an active area of 32 cm x 26 cm. Development of ionisation chamber arrays followed soon after with PTW (Freiburg, Germany) producing a 2D Array which initially had 256 vented ion chambers with cross-sections of 8 x 8 mm$^2$ spaced 1.6 cm centre-to-centre. This was then upgraded to the newer model seven29 which has 729 chambers with cross-sections of 5 x 5 mm$^2$ and spacing of 1 cm. The latter implementation is still in use; however it has recently been replaced commercially by the OCTAVIUS® 729 which has increased radiation shielding of the electronics.

These devices were originally developed to be able to measure per-beam fluence. In other words they were designed to be set up normal to the beam direction. At the time that detector arrays were becoming commercially available, electronic portal imaging devices (which had started being developed in the early 1980s) had matured with linac vendors, by 2001, using active matrix flat-panel arrays incorporating amorphous silicon (aSi) photodiodes [59]. These devices were designed to measure the MV beam transmission through the patient to be used for imaging and patient setup verification. However concurrently with ongoing imaging developments of EPIDs, it was recognised that with appropriate calibration they could also be used for dosimetry [60,61]. With sub-millimetre
resolution, they started to be used routinely for IMRT fluence verification. Therefore the use of detector arrays for these kinds of measurements became uncompetitive almost immediately and adaptations were made in hardware and software to be able to measured composite dose distributions. For the PTW 2D-ARRAY 729, this involved the development of the OCTAVIUS Phantom which was designed to reduce the angular dependence of the array at non-normal beam incidences [62]. For other planar detector arrays, cubic phantoms were used but developments in the calibration procedure were made such that the user is able perform measurements with their equipment to characterise the angular dependence which can then be corrected out of the composite measurement; this requires the use of an inclinometer to monitor the gantry angle for each beam delivery [63].

Other commercial detector arrays available are the Delta4® (ScandiDos AB, Uppsala, Sweden), ArcCHECK® (Sun Nuclear Corp., Melbourne, FL) and I’mRT MatriXX (IBA Dosimetry GmbH, Schwarzenbruck, Germany. Table 3.1 gives a summary of the main characteristics of commercial detector arrays that are currently available for IMRT and VMAT QA.
<table>
<thead>
<tr>
<th>Array configuration</th>
<th>PTW 2D-ARRAY 729 seven29 / OCTAVIUS® 729 Detector Array</th>
<th>PTW OCTAVIUS® 1500 §</th>
<th>PTW OCTAVIUS® 1000 SRS</th>
<th>IBA i’mRT MatrixXX®</th>
<th>Scandios Delta4®</th>
<th>Sun Nuclear MapCHECK® 2</th>
<th>Sun Nuclear ArcCHECK®</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detector type and shape</td>
<td>Plane-parallel vented ion chambers; cubic</td>
<td>Plane-parallel vented ion chambers</td>
<td>Liquid-filled ion chambers; cubic</td>
<td>Vented pixel ion chambers; disc shaped</td>
<td>p-type Diode; disc shaped</td>
<td>SunPoint® n-type Diodes; cubic</td>
<td>SunPoint® n-type Diodes; cubic</td>
</tr>
<tr>
<td>Detector size (mm)</td>
<td>5.0 × 5.0 × 5.0</td>
<td>4.4 × 4.4 × 3.0</td>
<td>2.3 × 2.3 × 0.5</td>
<td>4.5 diameter × 5 height</td>
<td>1.0 diameter × 0.05 height</td>
<td>0.8 × 0.8 × 0.03</td>
<td>0.8 × 0.8 × 0.03</td>
</tr>
<tr>
<td>Detector volume (cm³)</td>
<td>0.125</td>
<td>0.058</td>
<td>0.0026</td>
<td>0.08</td>
<td>0.000039</td>
<td>0.000019</td>
<td>0.000019</td>
</tr>
<tr>
<td>Detector spacing (centre-to-centre)</td>
<td>10 mm</td>
<td>7.1 mm</td>
<td>2.5 mm in central 5.5 x 5.5 cm and 5 mm elsewhere</td>
<td>7.62</td>
<td>5 mm in central 6 x 6cm; 10 mm to outer 20 x 20 cm</td>
<td>7.07 mm</td>
<td>10 mm</td>
</tr>
<tr>
<td>Maximum field size (cm)</td>
<td>27 × 27</td>
<td>27 × 27</td>
<td>10 × 10</td>
<td>24.4 × 24.4</td>
<td>20 ×</td>
<td>32 × 26</td>
<td>21 × 21</td>
</tr>
<tr>
<td>Number of detectors</td>
<td>729</td>
<td>1405</td>
<td>977</td>
<td>1020</td>
<td>1069</td>
<td>1527</td>
<td>1386</td>
</tr>
<tr>
<td>Housing material</td>
<td>PMMA / Glass-reinforced plastic</td>
<td>PMMA</td>
<td>Glass-reinforced plastic</td>
<td>Tecaran ABS buildup, RW3 backscatter</td>
<td>PMMA or Plastic Water®</td>
<td>PMMA (Acrylic)</td>
<td>PMMA (Acrylic)</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>3.2 / 5.7</td>
<td>6.0</td>
<td>5.4</td>
<td>10.0</td>
<td>24.0</td>
<td>7.1</td>
<td>15.4</td>
</tr>
<tr>
<td>Phantom shape for composite dose measurement</td>
<td>Octagonal (2D) Cylindrical (3D dose reconstruction)</td>
<td>Octagonal (2D) Cylindrical (3D dose reconstruction)</td>
<td>Cylindrical (3D dose reconstruction)</td>
<td>Cubic</td>
<td>Cylindrical</td>
<td>Cubic</td>
<td>Cylindrical (hollow central cavity) Separate CavityPlug™ available</td>
</tr>
</tbody>
</table>

* Note that the PTW 2D-ARRAY 729 seven29 has been discontinued commercially and replaced by the OCTAVIUS 729. It is included here as it has been used in this thesis.
§ Commercially available from October 2014
The PTW 2D-ARRAY 729, Delta4 and ArcCHECK were used in this thesis. Most of the experimental work in this thesis was on the PTW OCTAVIUS 2D-ARRAY 729 seven29 as this was the initial device available to the author and is featured in Chapters 4, 6 – 9. The Delta4 was owned by the Royal Surrey County Hospital and the ArcCHECK was kindly loaned by Imaging Equipment Ltd on behalf of Sun Nuclear Corporation for a limited time period. A fuller description of these devices is given below. These devices represent the three main device configurations as shown in Figure 3.2; these are planar (e.g. PTW 2D-ARRAY 729), helical (ArcCHECK) and cross-plane (Delta4). The PTW OCTAVIUS® 4D was recently developed which uses 2D measured data from any of the PTW detector arrays to reconstruct a 3D dose cube [64,65]. This, along with the OCTAVIUS 1000SRS detector array were loaned from PTW on a research collaboration agreement to investigate the differences in hardware resolution as described in Chapter 7.

3.2.1 Review of detector array characterisation studies for IMRT and VMAT QA

Various studies have previously been performed to assess the suitability of detector arrays for IMRT and VMAT QA. Letourneau et al [66] evaluated the dosimetric characteristics of the first commercial MapCHECK and found a linear dose response up to 2.95Gy, reproducibility within ±0.15%, calibration of the diodes to within ±1% of each other was achievable, and there was reported ‘good agreement’ with ion chamber and film results. Li et al [67] performed a comparison between the MapCHECK and the MatriXX for IMRT QA and reported ‘excellent passing rates’ for a set of 6MV and 18MV IMRT fields. Buonamici et al [68] performed an intercomparison between film dosimetry and the MapCHECK and reported the detection of deliberate deviations was as good as film, therefore their overall conclusion was that the MapCHECK was suitable as a replacement for film dosimetry in routine IMRT QA. Yan et al [63] compared the sensitivity of the MapCHECK against radiochromic
films for detecting deliberate MLC deviations and found that using a gamma index passing criteria of 2%/2mm showed the strongest sensitivity for detecting MLC changes and that the sensitivity of the MapCHECK was larger than radiochromic film. Masi et al [70] compared the 2D-ARRAY 729 with other systems of dosimetry, including Delta4 and MapCHECK, in the detectability of MLC positional deviations in 50 Elekta™ VMAT plans in different cancer sites and found high pass-rates for 3%/3mm gamma index, and suggested possibly moving to 3%/2mm for the arrays studied. Spezi et al. [71] and Poppe et al. [72] found that for step-and-shoot IMRT, 1mm MLC deviations could be detected for per-beam planar verification using the PTW 2D-ARRAY 729. Myers et al [73] performed a comparison between the PTW seven29 and the Delta4 for 15 clinical Tomotherapy QA plans and >90% gamma index passing rates for 3%/3mm for both systems. These were considered clinically acceptable; therefore the conclusion was that they were suitable as replacements for ion chamber and film dosimetry in routine QA. Syamkumar et al [74] characterised the response of the 2D-ARRAY 729 seven29 in 10 clinical RapidArc QA plans and using 3%/3mm gamma index criteria, concluded that it was suitable to use for routine QA. Chandraraj et al [75] compared EDR2 film, IBA I’mRT MatriXX, PTW seven29 and Delta4 for RapidArc and IMRT QA. They found that all 4 techniques yielded equivalent results with all achieving 3%/3mm passing criteria. The gamma index results of the 3 detector arrays were found to be within 5% of film [75]. Heilemann et al [76] used the PTW seven29 and Delta4 to assess the sensitivity of the gamma index to MLC misalignments in RapidArc QA and recommended that 2%/2mm should be used as passing criteria for these devices instead of 3%/3mm but stress that visual inspections should be made as 2%/2mm did not pick up all clinically significant errors. Van Esch et al [77] described a formalism for RapidArc clinical implementation and have evaluated the PTW seven29 and Delta4 as suitable dosimeters for routine patient QA. Zhu et al [78] compared the ion chamber, EPID, seven29, MatriXX and Delta4 in the verification of 12 VMAT plans of different sites and complexity. Ion chamber measurements were within 3% and the detector arrays passed the 3%/3mm criteria at >90%. Letourneau et al [79] evaluated the ArcCHECK for VMAT QA and found good passing rates for 3%/2mm criteria and that it was able to sufficiently detect small gantry rotation changes (up to 3°) and phantom setup errors of 1mm. It was also found to be suitable for evaluating individual control points in the VMAT delivery. Lin et al [80] also validated the suitability of using the ArcCHECK for VMAT verification; in this work, ArcCHECK measurements for plans with deliberate translational and rotational errors were compared against the TPS and an independent Monte Carlo model. Lang et al [81] made use of the MatriXX, Delta4, and ArcCHECK for pre-treatment QA of flattening filter free VMAT to assess dosimetric accuracy and Petoukhova [82] reported on HybridArc verification using the ArcCHECK. Bedford et al [83] performed benchmarking measurements for the introduction of the Delta4 into clinical use for IMRT and VMAT verification.
The study found that the Delta4 measured a dose within 2.5% of an ion chamber and a slightly higher passing rate for 3%/3mm gamma index than film. Feygelman et al [84] evaluated the Delta4 for Tomotherapy QA with mean passing rate of 97% for 3%/3mm in 9 clinical plans and recommended the use of MVCT imaging for phantom alignment. Fredh et al [85] performed a comparison of the Delta4, OCTAVIUS 729, COMPASS and Epiqa™ to assess their response to deliberate errors in VMAT patient-specific QA and found considerable variation in the types of errors that could be detected by the different systems, as well as poor correlation between the gamma index results and DVH deviations. Zhen et al. [86] performed a theoretical evaluation of three detector geometries: ‘X’, ‘O’, and spiral shapes which simulated modifications by modifying the beam models to introduce MLC transmission and penumbra errors, in order to create ‘virtual measurements’ at the treatment planning system resolution. These were compared against error-free calculations. This method meant uncertainties in delivery and devices were removed. A similar methodology was used by Nelms et al [87].

### 3.3 Description of Detector Arrays Used in This Thesis

Basic commissioning tests of the detector arrays were performed on a Varian Clinac iX (Varian Medical Systems, Palo Alto, CA). The Clinac incorporates the Millennium 120 leaf MLCs, with the central 80 MLCs covering 20x20cm each having a 0.5cm width at the isocentre; the remaining MLCs have 1cm width. The methodology used was in keeping with previously published reports [60,66,74,88–94]. All systems were used according to individual manufacturer recommendations. The tests performed for the basic commissioning of detector arrays are discussed in Appendix A, with the PTW 2D ARRAY used as an example.

#### 3.3.1 The PTW OCTAVIUS® Series

##### 3.3.1.1 VERISOFT®

For all of the different detector array systems offered by PTW, the software required is PTW VERISOFT® which is used to acquire and analyse measurements. Various versions became available throughout the duration of this PhD research. Versions 4.0 - 6.0 were therefore used at various times and the particular version is specified at the appropriate points. All versions shared the same gamma index implementation and mainly interface and upgrades to the handling of DICOM objects were the principal modifications between different versions.
3. Detector array technology

3.3.1.2 2D-ARRAY 729 seven29 / OCTAVIUS 729

The PTW 2D-Array consists of a matrix of 729 cubic vented ionization chambers with 0.5cm x 0.5cm cross-section, spaced 1cm centre-to-centre, giving a total area of 27cm x 27cm [89]. The upper electrode layer sits below a 0.5cm PMMA build-up layer whereas the lower electrode layer lies on top of a 0.2cm thick electrode plate which itself is mounted on a 1cm PMMA base plate. The nominal effective point of measurement (EPOM) is located at 0.75 cm from the surface. The OCTAVIUS phantom has an octagonal shape in its cross-section, and is designed to allow composite rotational IMRT plan verification. The phantom is made of polystyrene which has a physical density of 1.04g/cm$^3$. Its dimensions are 32 cm width, 32 cm length, 32cm height, and has a 30x30x2.2 cm$^3$ central cavity for the 2D-ARRAY 729 [62].

Figure 3.2 The PTW Ocavius II phantom with 2D-ARRAY 729 detector array in situ.

The OCTAVIUS phantom was CT scanned twice with both the 2D-ARRAY 729 in situ and with a homogeneous insert for comparison, see Figure 3.3. For composite field measurements, the base of the OCTAVIUS contained a semi-circular air gap to correct for the inherent under-response of the 2D-ARRAY 729 when the radiation field is incident posteriorly, as described by Van Esch et al [62]. For planning, the phantom was scanned with a solid base.
3. Detector array technology

3.3.1.3 **PTW OCTAVIUS Detector 1000\textsuperscript{SRS}**

The OCTAVIUS Detector 1000\textsuperscript{SRS} consists of a matrix of 977 liquid-filled ion chambers with 2.3 mm $\times$ 2.3 mm $\times$ 0.5 mm volume and 2.5 mm centre-to-centre detector spacing in the central 55 cm $\times$ 55 cm area and 5 mm in the outer 110 x 110 mm area. The nominal EPOM is located at 0.9 cm from the surface.

---

**Figure 3.3** CT scan of the OCTAVIUS phantom with *(left)* 2D-ARRAY 729 *in situ*, and *(right)* homogeneous insert.

**Figure 3.4** The PTW OCTAVIUS Detector 1000\textsuperscript{SRS}.
In a separate collaborative study [95], the $1000^{srs}$ array has been compared against Monte Carlo simulations, Gafchromic film, glass optical fibres, and glass beads in measuring small open field sizes down to $1 \times 1\text{cm}$ and showed excellent response in this study.

![Profiles acquired with the BEAMnrc/ DOSTYZnrc Monte Carlo simulation code (MC), Gafchromic film (GF), ionization chamber (IC), glass beads (GB) and optical fibre (OF) for the 10 × 10 cm, 4 × 4 cm, 3 × 3 cm, 2 × 2 cm and 1 × 1 cm field sizes at 5 cm water depth and normalized to the dose at the central axis for a 10 × 10 cm field size defined at the surface. From Jafari et al [95], with permission.](image)

### 3.3.1.4 PTW OCTAVIUS® 4D system

The OCTAVIUS 4D phantom is made of polystyrene. It is cylindrical and has dimensions 32 cm diameter and 34.3 cm length, and has a $30 \times 30 \times 2.2 \text{ cm}^3$ central cavity to accommodate the detector arrays. The phantom makes use of an inclinometer which is attached to the linac gantry to allow synchronous rotation of the OCTAVIUS 4D with the linac, see Figure 3.6. In this situation the detector array is always perpendicular to the radiation beam.
For predicted TPS dose calculation, a homogeneous cylinder of the same physical size as the phantom was used with a relative electron density in Eclipse set to 1.016 (~1.05 g/cm3). The required Verisoft software version was at least 5.1 and over in order to be able to perform 3D dose reconstruction; version 6.0 was used when measurements were carried out. Time integrated dose measurements at each gantry angle are acquired and the software uses the data to reconstruct a 3D dose distribution; the algorithm for this has been previously discussed [65,96]. To be able to perform the reconstruction, the software requires percentage depth dose (PDD) in water data for each beam energy used in the measurements for 85cm FSD and for square field sizes 2cm – 26cm. The PDD in water is converted to PDD in polystyrene (assuming relative electron density of 1.016). Generally speaking, for each gantry angle, the detector array measures the beam and calculates the effective field size using the detectors. Using the appropriate PDD for the effective field size, the dose plane is reconstructed at intervals along the ray line for that gantry angle; see schematic in Figure 3.7. This is done for all gantry angle measurements and summed together. Then the dose values are sorted into a 3D dose grid with a default spacing of 2.5 mm using linear interpolation [96].
3. Detector array technology

![Schematic representation of the 3D dose reconstruction algorithm used for the OCTAVIUS 4D. The red line indicates the measured dose plane. The blue lines indicate the reconstructed dose planes at intervals along the ray line for the gantry angle.](image)

3.3.2 Scandidos Delta4®

The Delta4® phantom consists of 1069 disc-shaped p-Si diodes, 1 mm diameter x 0.05 mm thick, arranged in a cross-plane configuration and housed in a PMMA cylindrical phantom. The dimensions of the phantom are 22 cm diameter and 40 cm length. The spacing of the detectors is 0.5 cm centre-to-centre in the central 6 cm x 6 cm area and 1 cm spacing in the outer 20 cm x 20 cm area. Measurements were performed with the inclinometer and synchronized to the linac trigger pulse as recommended [83]. A virtual CT cylindrical phantom with the same dimensions as the Delta4 was used according to the manufacturer specification. The software associated with this system carries the same name (Delta4 Software version: February 2012 release). Each detector board within the Delta4 had a relative uniformity calibration. An absolute dose calibration was performed against a Farmer-type ionization chamber. Both procedures followed the detailed instructions in the technical manual. The daily correction factor (DCF) procedure was used to normalise the output on the day of measurements as recommended [83].
3. Detector array technology

3.3.3 SunNuclear ArcCHECK®

The ArcCHECK® consists of 1386 n-Si diodes (0.8 x 0.8 mm) arranged in a helical shape at 3cm depth along the long-axis of a cylindrical phantom made of PMMA acrylic. The dimensions of the cylinder are 21 cm length and 21 cm diameter. The detectors are spaced 1cm centre-to-centre and measure an exit and entrance dose during delivery. All measurements were performed with 15 cm diameter CavityPlug™ homogeneous PMMA cylinder. A CT-scan of the ArcCHECK phantom was used for verification plan calculations. The scan was overridden where artefacts were caused by the diode detectors. The associated software used for acquisition and analysis was SunNuclear SNC Patient™ version 6.1. The methodology described in the technical manual was followed to calibrate the ArcCHECK. This involved; a background correction, a uniformity correction, an angular correction and an absolute dose calibration against a Farmer-type ionization chamber. The absolute calibration was done on the day of plan measurements.
3. Detector array technology

3.4 THE NEW TREND TOWARDS MEASUREMENT GUIDED 3D DOSE RECONSTRUCTION USING DETECTOR ARRAYS

There is a growing trend in measurement guided 3D dose reconstruction in patient anatomy. The basic premise is to use the dose measured by detector arrays to either back-project a dose onto the patient or to use the measurement to effectively ‘correct’ the TPS dose distribution according to the measured deviation. This then allows for a direct comparison between the reconstructed distribution and the TPS distribution in the patient anatomy and therefore DVH comparisons can be performed for volumes of interest. Recently, software algorithm and hardware improvements have led to the possibility of this technique. All of the main detector array manufacturers listed in Table 3.1 have focussed on implementing this technique and at the time of writing, all had introduced at least an early version of software for this kind of analysis.

3.4.1.1 The IBA Solution: COMPASS®

The MatriXX system can be adapted to be attached directly onto the Linac gantry head. The device has the addition of a digital inclinometer to monitor the gantry angle. This upgrade to the device has been branded MatriXX Evolution®. The device measures the actual beam fluence and then uses that to
calculate a dose distribution in the patient anatomy. This is achieved using the IBA COMPASS® software which has been developed in partnership with RayStation Laboratories (RaySearch Laboratories AB, Stockholm, Sweden). The software uses a collapsed cone convolution algorithm to calculate the dose (in fact as the software has its own algorithm it can be used as an independent dose check system of the TPS). Studies have been performed on evaluating this system; Boggula et al [97] evaluated the performance of the COMPASS system by comparing the reconstructed dose against Monte Carlo calculations and Godart et al [98] reported on the MLC error detection ability of the COMPASS software.

### 3.4.1.2 The PTW solution: DVH 4D

The PTW DVH 4D module relies on the use of the OCTAVIUS 4D’s 3D reconstruction dose algorithm which has been described in section 3.3.1.4. The CT data along with the RT structures are imported into Verisoft. The dose reconstruction in the CT anatomy follows the same principle as the dose reconstruction in the OCTAVIUS phantom, except that instead of assuming uniform density for the dose reconstruction along each ray line, the Hounsfield Units in the CT data are converted to relative electron densities [96]. The PDDs are converted to Tissue Phantom Ratios (TPR) according to the recommendations in BJR Supplement 11 [99]. For each detector along the ray line the water equivalent depth in the OCTAVIUS phantom is calculated; \( z_{\text{det}} \). For the CT voxels along the ray line, the water equivalent depth in the patient is calculated; \( z_{\text{CT}} \). Then the geometrical distance between the linac focus and the detector, \( a_{\text{det}} \) and the CT voxel \( a_{\text{CT}} \) respectively is determined. The dose at each voxel along the ray line, \( D_{\text{CT}} \) is then calculated as follows [96]:

\[
D_{\text{CT}} = D_{\text{det}} \left( \frac{\text{TPR}_{z_{\text{CT}}}}{\text{TPR}_{z_{\text{Det}}}} \right) \left( \frac{a_{\text{det}}}{a_{\text{CT}}} \right)^2
\]  

[3.1]

This algorithm only requires the CT data and therefore is independent of the TPS system; however the dose calculation is fairly basic and doesn’t take into account scatter effects. At the time of writing there have been no published reports on this as it was released in late 2014.

### 3.4.1.3 The Sun Nuclear solution: 3DVH®

This module works with MapCHECK for fixed field IMRT [86] or with ArcCHECK for IMRT or VMAT measurements. In both cases the measured dose is used to perturb the TPS dose distribution on the patient anatomy [100]. This method uses the Planned Dose Perturbation (PDP) [86] and ArcCHECK Planned Dose Perturbation (ACPDP) that has been described by Nelms et al [100]. In the ArcCHECK,
the author’s interpretation of this algorithm is as follows: firstly, time stamped entrance & exit dose measurements are used to discretise the treatment delivery into finite control points. Each control point is essentially the measured gantry angle using the ArcCHECK inclinometer as a function of time. These control points are used to modify the DICOM Plan file (which specifies the beam geometry, MU and planned control points for the MLC motion) to ‘synchronise’ the planned control points with what is actually measured; i.e. by modifying the planned gantry angle header with the measured value. The modified DICOM RT Plan is then processed by ACPDP which first generates a relative 3D dose distribution with a grid spacing of 2mm through fast Fourier transform (FFT) convolution of the time resolved control points, the total energy released per unit mass (TERMA) in the ArcCHECK and a 3D pencil beam dose kernel. For each sub-beam the x-ray beamlet is projected from the entry and exit surfaces of the ArcCHECK to achieve perturbation factors which are interpolated from entrance and exit values. The final absolute 3D dose distribution is then the 3D relative dose convolved with the scaling factors. Because of the way the perturbation factors are derived, only the TPS dose that lies within the dimensional space of the ArcCHECK can be perturbed; for regions outside this space, the software simply retains the original TPS dose value.

The disadvantage with this kind of implementation is that it relies heavily on information provided by the TPS and therefore there is an argument that it is not fully independent. Watanabe and Nakaguchi [101] evaluated the accuracy of the 3DVH module by comparing the 3D dose distribution reconstructed by the ArcCHECK with that measured using a polymer gel dosimeter. The gel was manufactured into a cylinder with similar dimensions to the ArcCHECK. In this study it was concluded that the 3DVH module produced an accurate 3D dose distribution when compared against the measured distribution in the gel dosimeter.

3.4.1.4 The Scandidos solution: Delta4DVHAnatomy®

The Delta4DVH® implementation is analogous to the ArcCHECK methodology, i.e. it is also a TPS dose perturbation algorithm. Hauri et al [102] performed an evaluation of the Delta4DVH Anatomy module for VMAT QA. However they found that the dose calculation algorithm was inferior to the TPS algorithm that they used in the study (Varian Eclipse AAA 8.9). This uses a pencil beam algorithm.
3. Detector array technology

3.4.1.6 EPID based systems for in vivo dosimetry

As previously mentioned, EPIDs have been in use for per-beam IMRT QA measurements since the early 2000s. In recent years there have been developments in performing 3D in vivo dosimetry using these systems [61]. The general principle is that by measuring the exit fluence through a patient using the EPID, it is possible to back-project that fluence and, through convolution with an appropriate energy deposition kernel, perform a 3D dose calculation in the patient CT data. This dose distribution can then be compared directly against the TPS distribution. Whilst the detector array system implementations highlight if there has been any error in the transfer of the treatment plan to the delivery system, EPID-based systems allow up to a day-to-day monitoring and additionally can be used to highlight any major changes to the patient anatomy that may have a clinical impact. At the time of writing, commercial systems for EPID dosimetry included Dosimetry Check™ by Maths Resolutions, Epiqa™ by EPIdos, and Sun Nuclear EPIDose™. However, only Dosimetry Check at the time was able to perform a 3D calculation in the patient and currently uses a pencil beam algorithm. The Netherlands Cancer Institute (NKI) in Amsterdam [103] and the MAASTRO Clinic in Maastricht [61,104] have performed extensive research on non-commercial systems. The NKI implementation uses a 3D back-projection algorithm and the MAASTRO implementation uses a Monte Carlo dose calculation algorithm to re-calculate the dose on patient anatomy. Strong commercial interest will likely lead to significant developments and competition in the next few years with systems utilising more accurate dose calculations, and with continuing advancements in Linac technology, the improvements in quality of on board image leading to the possibility of accurately calculating on the images will give rise to routine dose-guided-adaptive radiotherapy.

3.5 SUMMARY

Detector array technology offers the potential for novel approaches in complex radiotherapy QA. The technology continues to advance at the rate of advancement in the therapeutic systems. Various studies in the literature report that detector arrays are suitable for routine use in IMRT and VMAT QA. One aim in this thesis was to build on and add to that knowledge. At the time of the research work, some aspects of the above review had not been extensively looked at. For example, a comprehensive critical appraisal of the PTW 2D ARRAY 729 response to plan delivery deviations for VMAT had not been carried out; a methodology for this purpose was developed and published in the Journal of Applied Clinical Medical Physics [94], and is described in Chapter 4. A common theme in the review was the prevalent use of 3%/3mm gamma index passing criteria. There had been no
3. Detector array technology

study investigating whether this analysis was consistent between different commercial systems. Additionally there had been no studies looking at the combination of the software and hardware design effect on the consistency of the analysis. This research was carried out and published in Radiotherapy & Oncology [105], and is described further in Chapter 5 and 6. Additionally this work was extended to compare 2D planar gamma analysis and 3D volumetric analysis in Chapter 7, as well as the impact of the analysis on the resolution of detector arrays. At the time of this research, the majority of literature reports focussed on appraising detector arrays for routine QC. However, no study had investigated the suitability of a detector array in an audit setting. This was investigated as part of this thesis. The work is described in Chapter 8 and was published in Radiotherapy & Oncology [106]. Finally, alternative metrics to using the gamma index passing rate with detector arrays are investigated in Chapter 9, along with discussions about the methodology to establish appropriate acceptance thresholds.
4 DEVELOPMENT OF METHODOLOGY TO CHARACTERISE DETECTOR ARRAY TECHNOLOGY

4.1 INTRODUCTION

The purpose of this study was to develop a methodology to systematically characterize the response of a commercial detector array system for optimal use in composite clinical dynamic IMRT and VMAT verification. This study focussed on developing tests to evaluate the PTW OCTAVIUS II phantom and 2D-ARRAY 729 combination against the EPID and EBT2 Gafchromic film. The intention was then to use some of the tests developed to test the other commercial hardware and software combinations as discussed in chapter 6.

4.2 MATERIALS AND METHODS

4.2.1 EBT2 Gafchromic Film

Gafchromic EBT2 20 cm x 25.4 cm film sheets were used. Measurements were performed in a 30 cm x 30 cm x 20 cm solid water cubic slab phantom (Gammex Inc., Middleton, WI). The orientation of the film was consistent for all measurements; each film had a mark from manufacture to allow consistency in setup. All Gafchromic films were processed and analysed at least 24 hours after exposure. Films were scanned using the Epson Espression 10000 XL flatbed (Seiko Epson Corp., Nagano, Japan) colour scanner at a resolution of 75 dpi, using the red channel [91]. The films were all scanned in the same orientation and a jig was used to place the films in the same part of the scanner to minimize any displacement effects and to use the optimum part of the scanner which was determined through the commissioning process. A uniformity correction was applied by scanning a blank film from each batch. A calibration curve for the Gafchromic film batch was determined for a range of doses between 0 and 600 cGy. For analysis, the IBA OmniPro I’mRT v.7.0. software was used.
4. Methodology for detector array characterisation

4.2.2 Varian Electronic Portal Imaging Device (EPID)

The Varian EPID system used was the aSi1000 version which is a panel of amorphous silicon diode with a resolution of 0.392 mm. The measuring area of the EPID was 40cm x 30cm. Varian Portal Dosimetry software v.10 was used for analysis. In this software there was an option to create a composite of all individual fields. For consistency against the other systems this option was used for analysis. The EPID underwent a dark field correction (to correct for background noise), a flood field uniformity correction and an absolute calibration. The absolute calibration was done by delivering a 10x10 cm field size 6MV beam for 100 MU for a source-to-imager distance of 100 cm. The measured response was then set as 1 calibrated unit (CU) according to the definition in the Varian AM Maintenance imager calibration software.

4.2.3 Comparisons of using the OCTAVIUS scan with 2D-ARRAY 729 in situ vs homogeneous scan.

Calculating on the scan of the OCTAVIUS with the 2D-ARRAY 729 in situ with an advanced calculation algorithm may result in perturbation of the predicted dose by the air filled ionization chambers, which may add to uncertainties in dose, with impact when using gamma index analysis [45]. Therefore, a dosimetric comparison was performed between using the OCTAVIUS scan with the 2D-ARRAY 729 in situ and a homogeneous insert. All clinical plan composite measurements, described below in section 4.2.5, were evaluated using predicted doses calculated on both scans to compare the sensitivity of the gamma index analysis.

4.2.4 Multiple acquisition modes in the 2D-Array

In the PTW Verisoft software it is possible to merge multiple measurement acquisitions as proposed by Spezi et al [107]. The sequence of measurements is as follows:

1. a measurement is performed at the central axis, then
2. the 2D-Array is moved 0.5cm inferior, then
3. the 2D-Array is shifted 0.5cm to the right, then
4. the 2D-Array is shifted 0.5cm superior

By performing the above sequence and merging the measurements, it was possible to effectively increase the total number of measurement points four-fold from 729 to 2916, and improve the
4. Methodology for detector array characterisation

detector spacing from 1cm to 0.5cm centre-to-centre. For planar measurements, this can be easily achieved by automated couch movements. However, for composite measurements using the OCTAVIUS phantom, the 2D-Array must be shifted within the phantom and an insert is available to facilitate this. It may not be practical to perform this for every clinical plan verification, therefore the effect of different acquisition techniques was compared. All of the test fields and clinical plans described below were measured using the multiple acquisition technique. Comparisons were then performed between:

- a) Single acquisition
- b) Merging in the lateral direction only (by performing 2 acquisitions)
- c) Merging in the longitudinal direction only
- d) Full merge after four acquisitions

The gap between each ion chamber in the 2D-Array is 5mm wide as can be seen in the schematic in Figure 4.1. Suppose that only a single 5mm MLC leaf was being sampled. In this case three possibilities may occur for a collimator rotation of 0 degrees:

- i. Direct overlap between the leaf and a line of detectors
- ii. Partial overlap with a line of detectors
- iii. Complete miss if the leaf aligns with the gap between the lines of detectors

In normal situations, whereby the setup is such that the field’s cross-hairs align with the marks on the 2D-Array and OCTAVIUS phantom, the central axis will intersect the central detector. In this case, scenario 2 will occur and is illustrated in Figure 4.1 for static gantry IMRT where the collimator is typically set to 0 degrees and for RapidArc where the collimator angle may be typically set to 30 degrees. However scenario [i] and [iii] above would occur if a superior-inferior movement of 0.5cm is performed. In this case, every other MLC leaf will directly overlap with a row of detectors for a collimator angle of 0 degrees, and the remainder will be missed. This effect is minimised where there is a collimator rotation. Therefore in order to test the limits of the 2D-Array, comparisons were also performed using acquisition number 2 in the measurement sequence described above.
4. Methodology for detector array characterisation

4.2.5 Deliberate plan modification tests

The resolution and sensitivity of the 2D-ARRAY 729 was tested by a number of methods. All plans described in the following sub-sections were created using Eclipse™ and calculations were performed using the analytical anisotropic algorithm (AAA) with a 0.25cm grid spacing. The AAA is categorised as a 3D convolution-superposition dose calculation algorithm [108]. This type of algorithm computes dose as the superposition of the total energy released per unit mass (TERMA) with an energy deposition kernel which represents the spread of energy from the primary photon interaction site throughout the volume. The kernel is pre-calculated using Monte Carlo and in the case of the AAA, is a pencil beam type kernel. For further detailed information on the AAA, the interested reader is referred to the literature such as [108] and [109].

Measurements were performed on the same Varian Clinac iX over two sessions. Plans were generated to make optimal use of the 0.5cm MLCs. The array was cross-calibrated in the morning and afternoon of each session to account for any output fluctuation. In all cases the normal plan (i.e. with no modifications) was measured for baseline. Measurements were also performed using Gafchromic EBT2 film in the OCTAVIUS phantom in the same plane as the 2D-ARRAY 729, using the film insert provided with the phantom. In the case of the film measurements, the solid OCTAVIUS base was used and plans were calculated on a homogeneous scan. Gafchromic films were processed and analysed 24 hours after exposure. Films were scanned using the Epson Espression 10000 XL flatbed colour scanner at a resolution of 75 dpi, using the red channel [91].
In all cases, the gamma index (γ) method of evaluation was used with a 20% lower dose threshold [45]. Various criteria for γ were analysed, including the commonly used 3% dose difference and 3mm distance-to-agreement (DTA) criteria. For the 2D-ARRAY 729, analysis was performed using the PTW Verisoft software version 4.1. For film, analysis was performed using the Scanditronix Wellhöfer OmniPro™ I’mRT software version 1.7. Both the film and plan data were normalized at 100% to a point in a high-dose low-gradient region, to perform a relative comparison. This procedure is commonly used for film analysis due to the known difficulty in performing an absolute dose calibration for film [68]. In order to maintain a consistent comparison, the 2D-ARRAY 729 data was also re-scaled in the same way as the film. In both cases, the normalization point for the gamma analysis was kept consistent for any particular set of measurement; for example in the prostate IMRT plan with different changes introduced, the normalization value was always kept the same to avoid bias.

4.2.6 Gantry angle 0 degree test fields

As a starting point, it was necessary to understand the limitations of the 2D-Array in its basic IMRT measuring mode; that is setting the Gantry to 0 degrees and delivering a modulated field such that the 2D-Array is orthogonal to the beam. The aim was to investigate the two following questions:

- How does the detector spacing of the 2D-Array affect the measurement and visualization of a highly modulated field?
- How does the detector spacing affect the sensitivity of the gamma index in a modulated field?

Therefore, two individual planar test fields were designed specifically with the aim of addressing these questions. The test fields described below were also measured using the Varian aS1000 Electronic Portal Imaging Device (EPID) and Gafchromic film.

The first test field was designed to test the sensitivity of the gamma index analysis calculated in the 2D-Array using a modulated field with regions ranging from subtle to significant. This test will be referred to as the sensitivity test. The test had 54 regions of varying width and dose difference introduced into an open 15x15cm field using the fluence dose painting tool in Eclipse as shown in Figure 4.2 (left). The minimum spot size that the fluence painting tool allowed was 3mm width and 5mm height. As such, the columns in the test varied between one to six adjacent fluence spots (i.e. the width varied from 3mm in the first columns up to 15mm in the last column). Each row had a
height of 5mm and corresponded to a MLC leaf and the gap between each region was 5 mm. The difference in dose between each row and the high dose background ranged between 1% and 10%. The measurement of the field was compared to the open field predicted dose to determine the minimum detectable error by means of the gamma index analysis. Parameters for the analysis were varied from 1–10% dose difference, and 1–3mm DTA.

In the second test, a highly modulated field was created by dose painting varying dose and spatial positioning into an initially uniform field as shown in Figure 4.2 (right). This field is more complex than a field encountered clinically and tests the limits of the 2D-Array. The first two lines in the field were offset from each other by 5mm. This meant that due to the design of the 2D-Array, the resolution in the lateral direction could be tested. The third line increased in size in the longitudinal direction, and therefore the resolution could be tested in this direction. The remaining six lines were used as a combined spatial and dose resolution test. This test will be referred to as the resolution test. The field was delivered to the 2D-Array to test how well it performs in distinguishing the regions.

![Figure 4.2](image)

**Figure 4.2** Single gantry test fields for sensitivity (*left*) and resolution (*right*) assessment. In the sensitivity test the regions vary in width from left to right between 3mm and 15mm and the dose difference relative to the background varies from top to bottom by 1% to 10%. In the resolution test the values in the lower half represent difference in % dose between the regions and the background (lime green) area. In both tests, each row represents a single MLC leaf.

### 4.2.7 Test clinical plans with deliberate modifications

In order to evaluate the variability of the gamma index analysis in the different QA systems described above, changes were deliberately introduced to clinical treatment plans as has been used
by others [86,87,110,111]. Previous studies have simulated deliberate delivery errors by introducing MLC leaf bank changes or modifications to MLC transmission and penumbra [76,85,86,111]. In this study an alternative approach was employed. Single MLC leaf positional deviations of 1mm, 2mm, and 5mm were introduced, across the entire field, into pelvic and head & neck IMRT and RapidArc clinical plans by editing each control point within the RT Plan DICOM files. The plans which the MLC deviations were introduced into were: a 5-field prostate dynamic IMRT plan, a 6-field Head & Neck dynamic IMRT plan, a single 360° arc prostate RapidArc plan and a 2-Arc Head & Neck RapidArc plan. In the prostate plans the leaf chosen in all plans was such that the deliberate modification intersects the high-dose prostate region. In the head & neck plans the modification intersected the high dose region and the spinal cord sparing region. These were chosen as a relatively simple prostate plan and a complex head & neck plan to test how the arrays performed in those diverse scenarios. Deliberate collimator rotation changes of 1, 2 and 5 degrees were also introduced into a 2-Arc prostate & pelvic nodes RapidArc plan; this plan was chosen to give a large error as the area of the dose distribution covered most of the detector arrays. A further test was created using the fluence editing tool in Eclipse to manually dose paint hot and cold dose spots (ranging from -10% to +10% dose regions) of varying dimensions (0.5 cm x 0.5 cm to 2 cm x 2 cm) into all the fields in a 5-field prostate & pelvic node dynamic IMRT plan. These significant modifications were randomly introduced and were designed to be able to test both the resolution and response of the different systems. In this case a re-calculation of the leaf sequence was required. In total, 22 plans were created (5 of which had no modifications). This number of plans was practical to be able to perform measurements in each QA system in a few measurement sessions as described further below. To ensure that the fields were deliverable, all plans were re-calculated in Varian Eclipse v.10 using the analytical anisotropic algorithm (AAA).

4.2.8 Effect of normalisation point

The effect of choosing a point for the gamma index evaluation was investigated to assess whether this would influence the results. The analysis described above was repeated by deliberately choosing a dose point in a region where there was an MLC deviation, and by choosing the dose based on a mean value over the high dose region. In order to facilitate the latter, a custom spreadsheet was generated in Microsoft Excel 2007 (Redmond, Washington, US). The spreadsheet was created such that it was possible to import the 2D-ARRAY 729 measurement and predicted dose. The local % dose difference was calculated on a per detector basis, by comparing the measurement against the corresponding predicted dose. Customisable thresholds were also written into the spreadsheet such that the user may choose a lower and upper threshold for any value between 0% and 100%. The
maximum dose point was taken as 100%; assuming that the ICRU Report 83 [112] conditions were met for the high dose region, the coverage would range between 95% to 107% of the prescribed dose, i.e. a range of 12%. Allowing for changes in the homogeneity when the plan was transferred to the OCTAVIUS phantom, a lower threshold of 85% was used to ensure complete sampling of the primary PTV region. This spreadsheet was generated as it was found that the commercial systems (Verisoft and Omnipro) limited the lower threshold to a maximum of 30%, whether for dose difference or gamma analysis. It was then possible to acquire various statistics such as the mean of all the dose differences and standard deviation. The spreadsheet was also setup in a way that a comparison may be performed between one predicted dose plane and another.

4.2.9 Dosimetric and radiobiological evaluation of clinical plan modifications

The dosimetric impact of the subtle MLC positional modifications was assessed using the spreadsheet described in section 4.2.8. The predicted dose due to a MLC positional modification was exported to compare against the unperturbed predicted dose. The expected local dose difference caused by the deliberate MLC positional modification was calculated and compared to that found by the 2D-ARRAY 729. Additionally, the mean dose difference over a high dose region was also calculated.

In addition to the dosimetric impact, it was also possible to determine whether there is a theoretical radiobiological effect due to the changes introduced into the clinical plans as described by Carver et al [113]. Tumour control probability (TCP) calculations and normal tissue complication probabilities were performed in BIOPLAN [114]. TCP was calculated using the mechanistic Poisson-based TCP model [115]. The following input parameters for the TCP models for prostate tumours were used: radiosensitivity parameter $\alpha=0.29$ Gy$^{-1}$, inter-patient variation in radiosensitivity parameter $\sigma_\alpha=0.07$ Gy$^{-1}$, clonogenic cell density $\rho_c=10^7$ cm$^{-3}$ as well as an $\alpha/\beta$ ratio of 10 Gy [114]. For squamous cell carcinoma parameters were chosen as $\alpha/\beta=10$ Gy, $\alpha=0.305$ Gy$^{-1}$, $\sigma_\alpha=0.07$ Gy$^{-1}$, $\rho_c=10^7$ cm$^{-3}$ [116].

In the prostate cases, NTCP calculations were performed for the rectum using the Lyman-Kutcher-Bauman model [117–119], generalised uniform dose concept [120] and QUantitative Analysis of Normal Tissue Effects in the Clinic recommended best parameter estimates of $\alpha/\beta=3$ Gy, volume effects parameter $(n) =0.09$, slope parameter $(m) =0.13$, and the dose for 50% complication probability (TD50)=76.9 Gy. For the bladder, there is limited NTCP parameter data due to difficulties in fitting parameters to genitourinary toxicity [121]. The general consensus is to use the parameters...
of \( n=0.5, \ m=0.11, \ \text{and TD50}=80\ \text{Gy} \), in conjunction with \( \alpha/\beta=3\ \text{Gy} \) [122]. In the head & neck plans, NTCP values were calculated for spinal cord and parotids. For spinal cord, parameters for myelopathy were taken as \( \alpha/\beta=3\ \text{Gy}, \ n=0.05, \ m=0.175, \ \text{TD50}=66.5\ \text{Gy} \) [122]. For xerostomia \( \alpha/\beta=3\ \text{Gy}, \ n=0.7, \ m=0.18, \ \text{TD50}=46\ \text{Gy} \) [122].

4.2.10 Data and statistical analysis

To perform a quantitative analysis between the different permutations described above, a range of gamma index [45] passing criteria were recorded, including the commonly used 3%/3mm. For each passing criteria, the percentage of detectors with \( \gamma<1 \) was recorded. To compare the response of the different systems, cumulative histograms were plotted for the percentage of detectors/pixels passing with \( \gamma<1 \), for specified passing criteria, for a given plan; i.e. the number of plans where the percentage of points passing with \( \gamma<1 \) was 95%, 96%, 97% and so forth. It was possible to use the plots to illustrate the trend and agreement of the different systems against the independent predicted \( \gamma \). See the schematic in Figure 4.3.

![Schematic diagram of cumulative histogram analysis](image)

**Figure 4.3** Schematic diagram of cumulative histogram analysis for evaluating the gamma index response of different systems. On the y-axis is a plot of the percentage of all measured distributions and on the x-axis is plotted the percentage of measured points with \( \gamma<1 \). For a range of measured plans with no delivery errors one would expect all measured distributions to achieve 100% of measured points to pass with \( \gamma<1 \) for a given passing criteria. If a range of plans with known delivery ‘errors’ are measured, one would expect an ideal system to demonstrate a range of failures and produce a trend such as that illustrated by the solid blue line. As such the response of different systems can be tested by comparing their results against a baseline.
trend. Measured trends that tend to the left of the baseline have greater response and those to the right have a lower response. This type of graph allows for a comparison between different systems and can be used to compare different passing criteria.

In addition to the above analysis method, agreement between passing rates in the different measurement permutations and the 2D-ARRAY 729 measurement in single acquisition mode was statistically assessed using the concordance correlation coefficient, $\rho_c$ [123]. In the case of poor agreement, the statistical significance of any difference was assessed using the Wilcoxon signed rank test with $p<0.05$ as the threshold for significance.

### 4.3 RESULTS

#### 4.3.1 Gantry angle 0 degree test fields

Measurements of the resolution test field for the 2D-ARRAY 729 in single acquisition and full merge mode, Gafchromic film and EPID are shown in Figure 4.4. The array was able to distinguish dose differences, but there was a smoothing effect in the single acquisition which improved when the effective resolution was reduced to 5 mm.

![Figure 4.4](Image)

*Figure 4.4 (Left to right) Gafchromic film of highly modulated test, 2D-ARRAY 729 measurement (single acquisition), 2D-ARRAY 729 measurement (fully merged), EPID measurement.*

Passing rates for the sensitivity test using varying gamma index criteria are shown in Figure 4.5. The data have been plotted for the 2D-ARRAY 729 in single and full merge acquisition modes, film, EPID, and the expected passing rate. It can be seen that the DTA criteria had a minimal impact in this test field for all the permutations, except for the Gafchromic film. The single acquisition 2D-ARRAY data can be seen to be the least sensitive when compared to the expected passing rate. Spatial resolution was significantly affected; however dose resolution was less affected. This was due to the sparse resolution of 1cm. Improvements were found when a full merge acquisition was performed. As the dose difference criterion was increased, the different systems began to converge. The
Gafchromic film measurement, although very good spatially, appears to give false negative results when compared to the expected passing rate. This is due to intrinsic film heterogeneity causing minor artefacts combined with processing uncertainty, which were enough to disrupt the gamma index analysis passing rate, and are some of the known limitations of film dosimetry [124]. The EPID was found to be have the closest agreement to the predicted gamma index passing rate.

![Gamma index passing rates for the sensitivity test field for different measurement permutations. Points have been linked to provide a visual guide.](image)

4.3.2 Dosimetric and radiobiological impact of the deliberate clinical plan modifications

The ability of the 2D-ARRAY 729 to detect local dose differences caused by the MLC positional modifications is shown in Figure 4.6. There was a statistically good agreement between the dose difference detected by the 2D-Array and the expected difference ($\rho_c=0.96$). A 1mm MLC deviation caused up to a 1% local dose difference, whereas for a 2mm deviation this was between 1% - 3%, and for a 5mm MLC deviation, a local dose difference of between 3% – 6% was observed.

In the prostate IMRT plan, a 5mm MLC positional modification resulted in a 0.4% NTCP increase, whereas in the RapidArc plan, this was 1.2%. For a 2mm change, the increase in the IMRT and RapidArc plan was 0.3% and 0.9% respectively. In the prostate & nodes plan with collimator rotation...
changes, a 1 and 2 degree deviations resulted in an increased rectal NTCP of 3.0% and 3.2% respectively. In all prostate plans, bladder NTCPs were found to be 0%, although this may not be clinically relevant and is due to the difficulty of fitting parameters to genito-urinary toxicity [121]. In the head & neck plans, NTCP values for spinal cord did not increase and were 0.2% for all IMRT plans and 0.1% for all RapidArc plans; these values are in keeping with published data on the incidence of myelopathy at the 45 Gy level [125]. Similarly for the parotids, the maximum increase was limited to 0.2%. As expected, TCP values increased in all plans due to the increase in local dose from the MLC positional modifications. This was as high as an increase of 3% for a 5mm MLC modification and a collimator rotation deviation of 2 degrees.

Figure 4.6  Expected vs measured local dose difference due to the MLC positional modifications.

4.3.3  Composite verification of clinical plans

All the unperturbed plans had a γ<1 passing rate of 100% using 3%/3mm. When using 2%/2mm, the passing rate for all the plans was greater than 97%. The 5mm systematic MLC positional modifications were detected using 3%/3mm in the 2D-ARRAY 729 in the IMRT plans and in the prostate single-arc RapidArc plans. However, the 2mm systematic deviations were difficult to detect using 3%/3mm; the γ in the region where the deviations occurred was increased in comparison to the surrounding area but was still <1, and hence would not be detected as a fault; the deviation was detectable at 2%/2mm. For the head & neck 2-arc RapidArc plan, none of the MLC deviations were visible in the measurement and were also found to have a low impact in the expected gamma index maps. This is due to the plan having opposing collimator rotations on each arc to minimise the
tongue and groove effect and the errors may have been largely cancelled out. For the prostate & pelvic nodes RapidArc plan with collimator rotation changes, 3%/3mm gave a passing rate of >99% for deliberate 1 and 2 degree changes, and reduced to 92% in the presence of a 5 degree collimator angle deviation. The 1 degree change would have still passed at 2%/2mm with a passing rate of 99.3%. The 2 degree deviation, however, resulted in a passing rate of 94.1% and would have failed if a 95% threshold was used. At 2%/2mm, the passing rate for the 5 degree rotation plan was 74.6%. In the cases where the deliberate deviations were detectable using 2%/2mm, a passing criteria of 3%/2mm would have passed if a passing threshold of 95% was used; however, had a passing threshold of 98% been used then these measurements would have failed. Table 4.1 gives a summary of the average and minimum percentage of detectors/pixels passing with \( \gamma < 1 \) in all the plans, for 3%/3mm, 3%/2mm and 2%/2mm passing criteria for the different acquisition permutations.

The analysis of the effect of choosing a normalization point found that there was no significant difference between choosing a point in an unperturbed region, a point in a deviation region, or mean dose within the 85% isodose at 3%/3mm or 3%/2mm. At 2%/2mm there was a reduction in the passing rate in the analysis based on mean dose by 0.5% compared to the other two normalization techniques. This reduction was small but statistically significant (\( p<0.001 \)).

4.3.4 Comparison between expected gamma index passing rates in Verisoft and OmniPro

There was good statistical agreement between the expected gamma index passing rates calculated in Verisoft and OmniPro as indicated by the concordance correlation coefficient (\( \rho_c > 0.90 \) for all passing criteria). The average difference between the passing rates calculated by Verisoft and OmniPro was 0.5% and 1.1% for 3%/3mm and 2%/2mm criteria respectively. The difference was found to be statistically not significant (\( p>0.20 \) for all). It was therefore reasonable to use the average passing rate for the expected gamma index calculated by both software for each clinical plan to compare against that measured by 2D-ARRAY 729 and Gafchromic film.

4.3.5 Comparison between evaluations using CT scan with 2D-ARRAY 729 in situ vs homogeneous insert.

As shown in Figure 4.7 (a) and Table 4.1, there was a small difference between passing rates using a criteria of 3%/3mm, however at 2%/2mm using the scan with the 2D-ARRAY 729 in situ appeared to be more sensitive to deviations than comparing against the predicted dose calculated on the
homogeneous scan; where at 2%/2mm the average passing rate was 95.7% compared to 88.5% in the array scan.

Figure 4.7  
(a) Cumulative histogram of gamma index passing rates at 3%/3mm predicted doses calculated using 2D-ARRAY 729 in situ and homogeneous scan. (b) Cumulative histogram of multiple acquisition analysis calculated at 3%/3mm (solid lines) and 2%/2mm (dashed lines). (c) Cumulative histogram of 2D-ARRAY 729 versus Gafchromic film gamma index analysis calculated at 3%/3mm and 2%/2mm.
4. Methodology for detector array characterisation

Table 4.1 Summary of mean and minimum gamma index passing criteria for all various measurement permutations. A lower number indicates greater response to deviation detection. The concordance correlation coefficient, $\rho_c$, is also given assessing agreement with single 2D-Array acquisition.

<table>
<thead>
<tr>
<th>Device</th>
<th>Acquisition</th>
<th>3%/3mm</th>
<th>3%/2mm</th>
<th>2%/2mm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% detectors/pixels passing with $\gamma$ &lt; 1 and $\rho_c$</td>
<td>Mean</td>
<td>Min</td>
<td>$\rho_c$</td>
</tr>
<tr>
<td>2D-Array</td>
<td></td>
<td>Mean</td>
<td>Min</td>
<td>$\rho_c$</td>
</tr>
<tr>
<td>Single</td>
<td></td>
<td>98.9</td>
<td>92.0</td>
<td>-</td>
</tr>
<tr>
<td>Merged lateral</td>
<td></td>
<td>98.8</td>
<td>91.3</td>
<td>0.984</td>
</tr>
<tr>
<td>Merged longitudinal</td>
<td></td>
<td>98.6</td>
<td>88.8</td>
<td>0.948</td>
</tr>
<tr>
<td>Merged full</td>
<td></td>
<td>98.6</td>
<td>89.5</td>
<td>0.922</td>
</tr>
<tr>
<td>Merged full</td>
<td></td>
<td>98.6</td>
<td>89.5</td>
<td>0.922</td>
</tr>
<tr>
<td>Shift 5mm</td>
<td>98.5</td>
<td>85.7</td>
<td>0.875</td>
<td>97.8</td>
</tr>
<tr>
<td>longitudinal‡</td>
<td></td>
<td>99.0</td>
<td>91.8</td>
<td>0.851</td>
</tr>
<tr>
<td>Homogeneous scan</td>
<td></td>
<td>98.5</td>
<td>95.5</td>
<td>0.204</td>
</tr>
<tr>
<td>Gafchromic film</td>
<td></td>
<td>98.5</td>
<td>95.5</td>
<td>0.204</td>
</tr>
</tbody>
</table>

$\dagger$ acquisition (b) as described in section 4.2.4.
4. Methodology for detector array characterisation

4.3.6 Single vs multiple 2D-ARRAY 729 acquisition modes.

There was no significant difference in the gamma index passing rate at either 3%/3mm or 2%/2mm between performing a single or multiple acquisitions, as shown in the cumulative histogram in Figure 4.7 (b) and Table 4.1. It appears that performing a single acquisition is comparable to multiple acquisitions.

Figure 4.8 shows the gamma index maps and passing rates for the prostate IMRT plan with a 5mm single MLC leaf deliberate change. The first panel, Figure 4.8a, shows how the predicted gamma index distribution should look as a result of the deliberate change. The deliberate MLC change was systematic in all fields of the IMRT plan, however it should be noted that due to the variation of the modulation in the field, the variation of the leaf gap will result in a non-linear dosimetric impact. Therefore the band of raised gamma index values will not be homogeneous.

It can be seen that performing a merged lateral acquisition (Figure 4.8c) is visually comparable to a single acquisition (Figure 4.8b), whereas slightly improved resolution is achieved by either merging longitudinally (Figure 4.8d) or performing a full merge of four acquisitions (Figure 4.8e). The effect of the finite composition of the air filled ionisation chambers within the array can be seen in the measurements. It can also be seen in Figure 4.8 and Table 4.1 that the single acquisition was comparable to the acquisition shifted 5mm on the longitudinal axis (Figure 4.8f), demonstrating no reduction in response to delivery changes. Overall the acquisition shifted 5mm longitudinal was found to be the most responsive acquisition position based on the gamma index passing rates.
4. Methodology for detector array characterisation

Figure 4.8  Gamma maps using 3%/3mm criteria showing the effect of multiple acquisition modes for prostate IMRT plan with 5mm MLC positional modification. (a) predicted gamma index distribution; 99.5%, (b) Single acquisition; passing rate 96.8%, (c) two merged acquisitions with array shifted lateral for second acquisition; 97.0%, (d) two merged acquisitions with array shifted longitudinal for second acquisition; 96.2%, (e) four merged acquisitions to give effective 5mm resolution; 96.8%, (f) acquisition with 5mm shift in the longitudinal direction 96.1%.

4.3.7 2D-ARRAY 729 vs Gafchromic film.

Figure 4.9 shows a comparison between the gamma index distribution (using 3%/3mm) in the 2D-ARRAY 729, Gafchromic film and expected gamma index distribution for the head & neck IMRT plan with a 5mm MLC positional modification and prostate & nodes with randomly distributed changes. Regions of failure were comparable between the 2D array and Gafchromic film, with the array exhibiting the blurred effect due to its resolution. Neither system picked up all the modifications in the prostate & nodes plan with random changes.

Average and minimum gamma index passing rates using criteria of 3%/3mm were comparable for the 2D array and film as shown in Table 4.1 and the cumulative histogram in Figure 4.7(c). At 2%/2mm the 2D array appears to have resulted in a higher overall passing rate. For a passing rate of 85% or below the 2D array and Gafchromic films were comparable at 2%/2mm. For passing criteria
of 3%/3mm, all film planes achieved 95% passing rate or above, for the 2D array this was found to be 90.5% of measured planes. At 2%/2mm 33.3% of film planes achieved a passing rate of 95% or above, whereas for the 2D array it was 66.7%. Statistically, there was a poor agreement between 2D array and film as given by $\rho_c$ for each passing criteria. The difference between 2D array and Gafchromic film was statistically significant for passing criteria 3%/3mm or 3%/2mm ($p=0.048$ and 0.001 respectively), however it was not significant for 2%/2mm ($p=0.11$). When compared against the expected gamma passing rate, the 2D array result had a statistically more significant agreement ($\rho_c = 0.91$ for 3%/3mm, and 0.79 for 2%/2mm) than Gafchromic film ($\rho_c = 0.35$ for 3%/3mm, and 0.22 for 2%/2mm).

![Figure 4.9](image)

Figure 4.9  Comparison between gamma index distribution at a passing criteria of 3%/3mm for the head & neck IMRT plan with 5mm MLC positional modification (a) predicted, (b) 2D-ARRAY 729, (c) Gafchromic film and prostate & nodes plan with randomly distributed fluences modifications for (d) predicted, (e) 2D-ARRAY 729, (f) film.

### 4.4 DISCUSSION

For planar measurements of IMRT fields, the 2D-ARRAY 729 in single acquisition mode performed the worst in measuring the sensitivity and resolution test fields. Spatial resolution was significantly affected; however dose resolution was less affected. This was due to the sparse resolution of 1cm. Improvements were found when a full merge acquisition was performed. In measuring individual IMRT fields with the 2D-ARRAY 729 orthogonal to the beam, the resolution may be more influenced.
4. Methodology for detector array characterisation

by the modulated nature of the fields. This may have less significance in a prostate plan than in a head & neck cancer case. It appeared that the Gafchromic film measurement, although very good spatially, was giving false negative results in the sensitivity test. This is due to intrinsic film heterogeneity causing minor artefacts combined with processing uncertainty, which were enough to disrupt the gamma index analysis passing rate, and are some of the known limitations of film dosimetry [124]. The EPID was found to be the most effective of the different devices. In this regard, if using a 2D-Array for planar field measurement, it would be advisable to consider performing a full merge acquisition when measuring very complex planar fields.

In composite plan verification, the 2D-ARRAY 729 demonstrated good sensitivity to subtle MLC positional modifications. There was a reasonable comparison between the gamma index distributions generated by the 2D-Array and Gafchromic film. At 3%/3mm, passing rates were similar between the two systems. The 2D-Array did exhibit a higher passing rate at 2%/2mm compared to film. However, it was also interesting to see that the passing rates from the 2D-Array agreed better with the expected passing rates than Gafchromic film, consistent with the static gantry planar test fields. The effect of manually choosing a normalization point was found to be minimal but there was a statistically significant small difference between normalising based on a mean dose and a point in the measured distribution. Performing a normalization based on a mean dose would provide more consistency.

Performing a merged lateral acquisition was visually comparable to a single acquisition, whereas slightly improved resolution was achieved by either merging longitudinally or performing a full merge of four acquisitions. This is because, in the case of a lateral shift, resolution is only gained along the MLC leaf path, whereas merging in the longitudinal direction perpendicular to the MLCs allowed more sampling of the leaf bank. The single acquisition was also comparable to the acquisition shifted 5mm on the longitudinal axis, demonstrating no significant reduction in response to delivery deviations. The lack of difference between the different acquisition modes can be explained by the fact that on Varian linear accelerators, MLCs are arranged either side of the central axis. However, the 2D-ARRAY 729 is setup such that the central detector is aligned directly with the central axis. Therefore each chamber is always sampling two 5mm MLC leaves simultaneously. A 5mm offset in the longitudinal direction would result in every other MLC potentially being missed. It would therefore be recommended that if a longitudinal shift is required (e.g. for a long IMRT field where it is necessary to avoid irradiating the electronics) that the shift be made in whole centimetres. It also appears that calculating the expected dose on a homogeneous scan may be less
sensitive to errors than calculating on a scan with the 2D-Array in situ. This is due to underestimation of the dose from the lateral and oblique directions when using the homogeneous scan. There appears to be an international trend to use 3%/3mm with a 95% passing threshold. In this study, it was found that in terms of passing rate, the criteria of 3%/3mm masked deviations caused by deliberate collimator rotation changes of 1 and 2 degrees, as well as 2mm MLC positional modifications. The collimator rotation changes introduced in the prostate & nodes RapidArc plan caused the rectal NTCP to increase by about 3% which may be clinically significant. The 2mm positional modifications increased the rectal NTCP up to 0.9% in the prostate plans. These deviations were detectable using passing criteria of 2%/2mm with a 95% threshold or using passing criteria of 3%/2mm with a 98% passing threshold. For this system, these may be the recommended criteria to be used in order to detect deviations that may cause a clinically significant increase in NTCP. The changes introduced all increased local dose difference and therefore the TCP was increased. One limitation of this study would be that none of the plan modifications resulted in a reduction of TCP. All the MLC positional modifications were designed to increase the leaf gap. Errors with closed leaf gaps were not created as there was a risk of causing MLC collisions. MLC positional modifications with narrower leaf gaps would have been expected to cause dose reductions. It was shown in Figure 4.6 that the 2D-ARRAY 729 was able to detect the dose differences caused by the MLC positional modifications and the strong linear relationship between the expected dose difference and the measured difference suggests that dose reductions may have been detected equally. It is suggested that the gamma index passing thresholds be used for guidance, but also be combined with a visual inspection of the gamma index distribution and calculation of the dose difference to assess whether there may be a clinical impact in failed regions.

4.5 CONCLUSIONS

Tests have been employed to characterise the sensitivity and resolution of the PTW 2D-ARRAY 729 and OCTAVIUS II phantom combination. The 2D-Array in single acquisition mode was comparable to multiple acquisition modes and Gafchromic film for composite IMRT and RapidArc plan verification. A gamma index criterion of 3%/3mm may potentially mask clinically relevant deviations. A criterion of 3%/2mm with a passing threshold of 98% or 2%/2mm with a passing threshold of 95% was found to be more sensitive in conjunction with an evaluation of the gamma index distribution. These tests have resulted in an understanding of the 2D-Array’s limitations and increased confidence in its use for clinical IMRT and RapidArc verification.
5  
EVALUATION OF THE IMPACT OF THE GAMMA INDEX CALCULATION APPROACH: A BESPOKE MATLAB SOFTWARE

5.1 INTRODUCTION

As described in Chapter 2, the gamma index is a computationally expensive process due to the need to search all points in the evaluated distribution. This becomes more complex when comparing two 3D dose distributions. Ideally the gamma index would be calculated quickly to give a result within a reasonable time. The computer hardware used will have an impact on the speed of the γ calculation. Given previous limitations with computer technology, a number of studies in the literature have focussed on ways to mathematically decrease the calculation time of the gamma index.

In order to investigate the different approaches that can be used to calculate the gamma index, a software tool was written and implemented in Matlab v2012a – 2014a (Mathworks Inc.). Some open source tools are available online, however these are limited. Matlab uses some in-built functions for manipulation of matrices that can simplify the coding of the gamma index algorithm, but require the two compared datasets to have the same matrix size. This new software was written so that it could accept two datasets with different resolution and matrix sizes, and by default it was set up to perform no interpolations on the reference dataset. This tool was implemented as a graphical user interface (GUI) for user-friendliness as shown in Figure 5.1.
5. Bespoke gamma index software

Figure 5.1  Screenshot of the gamma index calculation software implemented in Matlab.

The current features of the software are:

- Able to handle 2D or 3D DICOM dose distributions, Excel, or PTW 2D-ARRAY 729 measurement format
- Datasets do not have to have the same resolution/matrix size as each other
- For 3D DICOM, can visualise axial, coronal, or sagittal viewing plane
- Display dose profiles
- Perform 2D plane vs 2D plane γ analysis
- 2D plane vs 3D volume γ analysis
- 3D volume vs 3D volume γ analysis
- Specify whether to search the whole evaluated distribution, or limit the search to a user-defined distance from each reference point
- Perform global or local γ calculation, with ability to set dose difference and distance criterion.
- Allows the user to set lower and upper dose thresholds (a yellow outline is given to visualise this; can be used e.g. to focus on a high dose region)
- Allows the user to specify interpolation factors for either dataset 1 or dataset 2 and also to specify the type of interpolation algorithm; linear or cubic spline.
5. Bespoke gamma index software

- Allows the calculation of different gamma index metrics; % of points passing with $\gamma<1$, mean $\gamma$, median $\gamma$, maximum $\gamma$, or the minimum gamma index in the top X% pixels (e.g. minimum $\gamma$ in the top 1% pixels; $\gamma_{1\%}$).

- Calculate mean, median, and standard deviation dose difference within the user defined threshold.

- Toggle between displaying gamma index dose distribution, gamma index histogram, or dose difference distribution.

- Batch analyse PTW 2D-ARRAY 729 measurements

By default, the dataset 1 was designated as the evaluated distribution, and dataset 2 was designated as the reference distribution (according to the definitions specified in Chapter 2).

5.2 Optimising the Software

5.2.1 Impact of Limiting the Search Distance on Calculation Time

A global 3%/3mm gamma index comparison of 2D matrices both with 81 x 81 points at 2.5mm spacing took ~380s to complete on a PC desktop with a quad-core Intel i7 4GHz CPU, and 16GB of RAM. One simple way to speed up the $\gamma$ calculation significantly is to limit the search in the evaluated dose distribution to a certain distance around each reference point. An interesting observation made in the study by Wendling et al [52] is that by setting a limited maximum search distance, it is only necessary to calculate the distance between a reference point and all the evaluated points bound by the search distance once; this can be defined as a 2D or 3D array, $R$. Similarly, it would then be possible to calculate the dose difference between the reference point and all the evaluated points which can be defined as an array $D$ which has the same size as $R$. This significantly reduces computing overhead and makes it possible to perform matrix operations which are optimised for speed in programmes such as Matlab by calculating all elements using parallelisation. Equations 3.2 and 3.3 can then be used to calculate the gamma index for the reference point.

To evaluate the suitability of using a limited search distance, a two 360° arc Head & Neck RapidArc treatment plan was calculated using the Varian Eclipse Treatment Planning System (TPS) using the Analytical Anisotropic Algorithm (AAA) v11 for a dose grid of 2.5mm. The head & neck RapidArc plan was then copied and a collimator angle change of +5 degrees was introduced to the two Arcs. This changed plan was compared against the normal plan using the software. The changed and normal plans were exported in DICOM format to be evaluated against each other in the software. The
reason for the choice of a large collimator change was that this would introduce a range of failed and passing points with varying levels of dose gradient as shown in Figure 5.2.

For global gamma index passing criteria 3%/3mm, 3%/2mm, 3%/1mm and 2%/2mm, the search distance was linearly increased starting from no limitation. In this case the most informative parameter is the maximum calculated gamma index value as this will be affected by limited search distances. The calculation time was also measured for each permutation. As the distance criterion is varied, the physical search distance will have a different impact depending on the number chosen for $\delta r$. Therefore rather than plotting the maximum gamma index against search distance it was more meaningful to plot against the ratio of the search distance divided by $\delta r$ as shown in Figure 5.3. In this graph it can be seen that for different passing criteria there was a consistent trend towards the maximum gamma index having no variation once the ratio of search distance / $\delta r$ became $\geq 1.5$. The maximum gamma index was the same above the threshold as that where the entire evaluated distribution was searched. Even up to a higher ratio of 5 the calculation time was still significantly small at $\sim 0.4s$. Given these results, the search distance was by default set to be $3 \times \delta r$ taking into account the use of the software for 3D data.
5.2.2 Data Interpolation techniques

A fundamental issue with the gamma index calculation is that the result will be influenced by the data point spacing of the evaluated dose distribution. There can be inaccuracies when the pixel spacing is $\approx$ distance criterion and can lead to overestimation of the gamma index. Previously Low et. al. \cite{45} recommended that the pixel spacing should be $\frac{1}{3} \delta r$. More recently, Wendling et. al. concluded that the spacing should be of the order of $\frac{1}{10} \delta r$ \cite{52}. In order to achieve a smaller point spacing, it is necessary to interpolate the evaluated dose distribution to a finer grid size.

In the software, the in-built 2D and 3D interpolation functions (called interp2 and interp3 respectively) in Matlab were used. The software was designed so that it is possible to interpolate either dataset 1 or 2 for investigative purposes (dataset 2 was always defaulted to have no interpolation unless the user specified otherwise). The interpolation options available were to use (1) a linear interpolation, (2) nearest neighbour, (3) cubic method or (4) spline method. In order to test the different interpolation techniques available, the head & neck RapidArc treatment plan was calculated in the Eclipse TPS using 1.25mm, 2.5mm and 5mm grid spacing. Each individual grid size calculation was exported in DICOM format. The 2.5mm and 5mm grid calculations were then
interpolated to a grid size of 1.25mm grid size using the different interpolation techniques and compared directly against the 1.25mm dose calculation using global gamma index and 1% / 1mm passing criteria. The % points passing with γ<1, and mean gamma index were quantified as well as visual inspection of the gamma map for each interpolation technique. No lower threshold was used for the γ calculation.

Figure 5.4 shows the calculated gamma index maps for the four different interpolation techniques and Table 5.1 shows the % points passing γ<1 and mean gamma index results. It was found that the spline algorithm gave the closest agreement. Errors mainly occurred at the penumbral edges which have been highlighted using 1%/1mm criteria and is therefore within acceptable uncertainty. For the purposes of this software the spline algorithm was therefore deemed to be appropriate.

Figure 5.4 Comparison of different Matlab interpolation algorithms; (a) spline, (b) linear, (c) cubic and (d) nearest neighbour.
Table 5.1 Comparison of different interpolation techniques using global 1%/1mm, 0% threshold.

<table>
<thead>
<tr>
<th>Interpolation type</th>
<th>Passing rate (%)</th>
<th>Mean gamma index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spline</td>
<td>96.6</td>
<td>0.25</td>
</tr>
<tr>
<td>Linear</td>
<td>94.4</td>
<td>0.30</td>
</tr>
<tr>
<td>Cubic</td>
<td>96.4</td>
<td>0.26</td>
</tr>
<tr>
<td>Nearest neighbour</td>
<td>83.4</td>
<td>0.47</td>
</tr>
</tbody>
</table>

After confirming the spline algorithm for interpolation, it was necessary to benchmark how much interpolation is needed. The Head & Neck RapidArc plan used previously was employed again. The normal plan was again set as the evaluated distribution and the gamma index was calculated by varying the level of interpolation. In Matlab an integer interpolation factor, $I$, is specified. This changes the original spacing, $x_0$, by:

$$x_0 \over 2^I$$ (5.1)

For passing criteria of 3%/1mm, 3%/2mm, 3%/3mm, $I$ was varied from 0 to 7 in increments of 1. For each passing criteria the following ratio was calculated:

$$\frac{\delta r}{x_0}$$ (5.2)

For each calculation, the mean gamma index was calculated and the time taken for the calculation was recorded. The results of this are shown in Figure 5.5 where it can clearly be seen that the mean gamma varies sharply when the ratio of $\delta r$ and $x_0$ are close to 1 and begins to stabilise by the time the ratio has become 10. This is in keeping with the recommendations by Wendling et al [51]. Up to this ratio, the calculation time was only of the order of 1 second and therefore as the default for further evaluations using the software the interpolation factor was set so that the ratio of equation 5.2 was ≥10.
5. Bespoke gamma index software

Figure 5.5  Comparison of the maximum gamma index as a function of varying the interpolation factor for global gamma passing criteria 3%/3mm, 3%/2mm, 3%/1mm, and 2%/2mm. The calculation time taken is plotted against the right axis for each passing criteria.

5.3 COMPARISON AGAINST TWO COMMERCIAL SOFTWARE

The bespoke Matlab software was tested against two other established commercially available software; namely the PTW Verisoft v5.1 package and IBA OmniPro v7.0. The same combination of the normal and changed head & neck RapidArc plan was used. In order to test the software, the changed plan was calculated and exported as 2.5mm, 5mm and 10mm dose grids to compare against the 2.5mm grid normal plan. Global gamma index comparisons were made using 3%/2mm passing criteria and no lower dose threshold. The mean gamma and % passing rates were compared. Figure 5.6 shows gamma index maps for the 10mm grid spacing changed plan against the normal plan.

The top two images are from OmniPro and Verisoft respectively. It is clear that the OmniPro software prefers to plot the gamma index map as an intensity image where each pixel is given a discrete colour based on the gamma index. The Verisoft software uses a colour contour approach which is visually easier to interpret. The two styles were replicated in Matlab and the maps are given
5. Bespoke gamma index software

for the bespoke software below the respective commercial software maps. The figure shows good visual agreement in the gamma index maps between the bespoke software and the commercial packages.

![Gamma Index Maps](image)

**Figure 5.6** Gamma index map for the 10mm resolution reference distribution from (a) OmniPro, (b) Verisoft, (c) Matlab plotted using the *imagesc* function, and (d) Matlab plotted using the *contourf* function.

The comparison between the γ passing rate in the Matlab software, OmniPro and Verisoft is shown in Figure 5.7. This shows very good agreement between the Matlab and Verisoft calculations for all the different grid spacing of the deliberately changed plan.
Interestingly, if the Matlab software is re-run with no interpolation setting, there is excellent agreement against OmniPro. Clearly this shows evidence that Verisoft interpolates the evaluated dose distribution into finer grid spacing, whereas OmniPro does not. A similar trend was found with the mean gamma index shown in Figure 5.8. This is a good example of how there can be variability in the implementation of the gamma index calculation. This brief study focussed purely on the software side, however most commercial systems are designed taking into account the associated detector array configuration and therefore a more robust analysis will include an evaluation of the
combined hardware and software. This topic is addressed further in chapters 6 and 7 of this thesis. An extract of the code is provided in Appendix C, the entire code for the GUI consisted of ~4600 lines.

Figure 5.8  Comparison of mean gamma index between (top graph) Matlab, Verisoft and OmniPro analysis for reference distributions with 2.5mm, 5mm and 10mm pixel spacing, and (bottom graph) repeated using Matlab with no interpolation. Analysis used global 3%/2mm with no lower dose threshold.
6

A COMPARISON OF THE GAMMA INDEX ANALYSIS IN VARIOUS COMMERCIAL IMRT/VMAT QA SYSTEMS

6.1 INTRODUCTION

Previous studies have been performed to inter-compare different commercial detector arrays and the response of γ IMRT and/or VMAT simulated errors. Zhen et al. [86] performed a theoretical evaluation of three detector geometries: ‘X’, ‘O’, and spiral shapes which simulated errors by modifying the beam models to introduce MLC transmission and penumbra errors, in order to create ‘virtual measurements’ at the treatment planning system resolution. These were compared against error-free calculations. This method meant uncertainties in delivery and devices were removed. A similar methodology was used by Nelms et al. [87]. However, detector arrays are limited by their sparse spatial resolution which may affect the response of the gamma index analysis due to under-sampling [126].

It is, therefore, of interest to understand the variability in gamma index analysis between different commercial QA systems, including their shape, configuration and detector resolution, and their associated software. Hence, the purpose of this study was to compare the gamma index analysis in the commercial 2D and 3D detector arrays to assess the impact of low resolution in combination with the gamma calculation implementation. This study made use of the methodology developed in Chapter 4 to evaluate the variability of the resulting gamma index assessment across the systems. In keeping with the methodology proposed by Nelms et al. [87] and Zhen et al. [86], the predicted gamma index was calculated by comparing high resolution calculation of the deliberate errors in the different commercial 2D and 3D detector configurations against the error-free calculation, in each respective commercial software. The predicted calculation in each software was also compared against the calculation in the Independent Matlab Software (Chapter 5). The final goal of this study was to compare the gamma index calculated based on experimental measurements in the commercial systems against the predicted gamma index.
6. Comparison of the gamma index in commercial systems

6.2 Detector Array Systems Inter-compared

The commercial detector arrays used in this study were the PTW 2D-Array in the OCTAVIUS II phantom, SunNuclear ArcCHECK, and Scandidos Delta4. These are representative of each of the currently available detector array configurations which are: planar (2D-Array), cross-plane (Delta4) and helical (ArcCHECK). In addition, in-phantom EBT2 Gafchromic film was used as well as the Varian Electronic Portal Imaging Device (EPID) to provide high resolution measurements to experimentally compare against the potential under-sampling effects from using sparse detector arrays.

6.3 Virtual Plan Measurements to Calculate Predicted \( \gamma \) in Commercial Software

The influence of the hardware and linac delivery on the \( \gamma \) calculation was removed using a similar methodology to that employed by Nelms et al. [87] and Zhen et al. [86], which allows for a theoretical direct comparison between different commercial software as well as different array configurations and shapes. To perform this analysis, all the test plans developed in Chapter 4 with the deliberately inserted modifications, were calculated on the respective phantom CT scan in Eclipse, in addition to the original unperturbed plans. This effectively meant that it was possible to simulate the predicted gamma index pass rate in ‘ideal’ conditions as it was possible to remove any inherent QA hardware uncertainty, any inherent mechanical effects on the other unperturbed MLC leaves and no output fluctuations would be present. Additionally, this method removes the effect of under-sampling and/or blurring effects that are inherent in array designs.

The original dose distribution, in DICOM RT Dose format, was then imported into each of, Verisoft, OmniPro I’mRT, and SNC Patient and Portal Dosimetry, as the ‘reference’ data set and the perturbed dose was imported as ‘measurement’ dataset. It was then possible to calculate a predicted \( \gamma \). This was not possible in the Delta4 software as there is no straightforward way to replace the measured dose with virtual data. Each system was used to process the predicted doses as they normally would in the presence of a measurement. For example in the ArcCHECK the dose cylinder is unrolled to compare against the unrolled measured dose [92]. For Verisoft and OmniPro, coronal planes corresponding to the position of the 2D-Array or Gafchromic film were chosen for this analysis. For Portal Dosimetry, a two-dimensional predicted dose was calculated using the algorithm within Eclipse, and the predicted plane with no modification was compared against the predicted plane from the modified treatment plan.
To perform a consistent evaluation, an independent gamma index calculation using a code developed in MATLAB [127]⁴, was also performed. This was a 2D calculation. For independence, the plans were calculated on a separate virtual phantom which was a water equivalent cylinder with 30cm diameter and 30cm length, and a coronal plane through the axis of the phantom was used. The purpose of the independent calculation was to remove the impact of software-specific gamma calculations.

For the Delta4 software, it was not possible to directly inter-compare two dose cubes against one another. However, it was possible to compare the normal plan measurement against the deliberately changed plan. Since the measurements were carried out sequentially, any inherent systematic uncertainties will be present in both the normal and changed plan measurements and consequent differences will be minimised. Whilst this was not an ideal comparison, it allowed the software to be evaluated. To ensure the validity of this, a similar comparison was made using the 2D-Array data and compared with the results from the Delta4 and the independent gamma index calculation.

6.4 EXPERIMENTAL PLAN MEASUREMENTS

Following on from the theoretical derivation of the predicted gamma index, experimental measurements on the linac using the QA systems were performed. All test plans were delivered using the same monitor units (MU) as the original plan and compared against the original unedited plan. In order to ensure that there would be no problem when delivering the deliberately introduced changes, each plan was first delivered to the EPID and compared with the calculated version of that plan. The same 22 plans were measured by each of the five QA systems. All measurements were undertaken on the same Varian Clinac iX. The linac incorporates the Millennium 120 leaf MLCs, with the central 80 MLCs covering a 20x20cm area each having a 0.5cm width at the isocentre; the remaining MLCs have 1cm width. Due to the number of plans, it was not possible to perform measurements by all the different QA systems on the same day. However, measurements were performed on sequential days and routine dynamic MLC quality control (QC) checks using the EPID were performed each day, to minimize uncertainty. The routine tests include the picket fence,

⁴ This study was performed before the author fully developed a gamma index Matlab code. In the interest of time and in order to publish in the high impact journal, *Radiotherapy & Oncology* [105], a fruitful collaboration was set up with Dr Pejman Rowshanfarzad and Professor Martin Ebert (University of Western Australia) who, after being sent the DICOM files, performed the independent calculations using their own already developed Matlab code.
6. Comparison of the gamma index in commercial systems

sweeping gap, and MLC speed test, which have been described in the literature for IMRT and RapidArc [128–130]. Measurement of the QC tests were compared directly against a baseline measurement taken at linac commissioning using gamma index with strict criteria of 2% dose difference (DD) and 1mm distance-to-agreement (DTA) criteria and 95% passing threshold.

6.5 DATA AND STATISTICAL ANALYSIS

In all theoretical and experimental cases, the global γ were calculated with a 20% threshold relative to a point selected in the centre of the high dose region, with low dose gradient. For 2D-Array, Delta4, ArcCHECK and film measurements, the gamma index evaluation was performed taking into account the 3D dose distribution; i.e. the complete TPS dose cube, not just the dose distribution of the measured plane. Various γ criteria were analysed, including the commonly used 3%/3mm.

In addition to the cumulative histograms, the statistical agreement between the predicted γ calculated in the different systems was assessed using the concordance correlation coefficient, ρc [123]. The Pearson correlation coefficient assesses whether there is a linear trend between two datasets, but does not indicate whether they agree. However, ρc assesses the correlation between a trend on a scatter plot against the 1:1 trend expected if two measured datasets agreed. The agreement between measured γ and the predicted γ was also evaluated using ρc.

6.6 RESULTS

6.6.1 Predicted γ based on virtual measurements

There was statistically good agreement between the predicted γ from each software and the independent calculation (all ρc>0.92), also shown in the scatter plot in Figure 6.1 and the cumulative histogram plotted in Figure 6.2. A summary of the predicted mean and minimum percentage of detectors/pixels passing with γ<1 is given in Table 6.1 for 3%/3mm, 3%/2mm and 2%/2mm for each system. The plot for the Delta4 is, as described in the methodology section, given for the comparison between the experimental measured plan with and without error as it was not possible to compare two calculated dose distributions. This was shown to be valid as this approach tested with the PTW 2D-Array measured data gave ρc against the independent calculation of 0.96, whereas for Delta4 this was 0.93, for passing criteria 2%/2mm. This is further supported as the corresponding PTW VeriSoft virtual measurement result was 0.95.
### Table 6.1
Summary of the mean and minimum measured gamma index passing criteria for each system. The concordance correlation coefficient, $\rho_c$, is also given assessing agreement with independent gamma index. The software are listed in the same order as the associated measurement system.

<table>
<thead>
<tr>
<th>System</th>
<th>% detectors/pixels passing with $\gamma &lt; 1$ and $\rho_c$</th>
<th>3%/3mm</th>
<th>3%/2mm</th>
<th>2%/2mm</th>
<th>Mean</th>
<th>Min</th>
<th>$\rho_c$</th>
<th>Mean</th>
<th>Min</th>
<th>$\rho_c$</th>
<th>Mean</th>
<th>Min</th>
<th>$\rho_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Software predicted</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verisoft v5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>99.0</td>
<td>89.9</td>
<td>0.97</td>
<td>98.4</td>
<td>83.9</td>
<td>0.95</td>
<td>97.2</td>
<td>75.4</td>
<td>0.95</td>
</tr>
<tr>
<td>SNC Patient v6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>98.7</td>
<td>84.5</td>
<td>0.97</td>
<td>98.0</td>
<td>78.5</td>
<td>0.99</td>
<td>96.4</td>
<td>70.0</td>
<td>0.96</td>
</tr>
<tr>
<td>Delta4 software</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>98.8</td>
<td>89.4</td>
<td>0.96</td>
<td>98.3</td>
<td>84.9</td>
<td>0.93</td>
<td>97.3</td>
<td>77.3</td>
<td>0.93</td>
</tr>
<tr>
<td>OmniPro I’MRT v7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>98.7</td>
<td>82.6</td>
<td>0.95</td>
<td>97.9</td>
<td>73.8</td>
<td>0.97</td>
<td>96.2</td>
<td>57.1</td>
<td>0.92</td>
</tr>
<tr>
<td>Portal Dosimetry v10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>98.7</td>
<td>84.7</td>
<td>0.97</td>
<td>98.0</td>
<td>73.6</td>
<td>0.96</td>
<td>97.5</td>
<td>68.2</td>
<td>0.92</td>
</tr>
<tr>
<td><strong>Independent predicted</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>98.8</td>
<td>87.0</td>
<td>-</td>
<td>97.9</td>
<td>78.0</td>
<td>-</td>
<td>96.4</td>
<td>58.1</td>
<td>-</td>
</tr>
<tr>
<td>Measured</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTW 2D-Array</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>98.0</td>
<td>86.3</td>
<td>0.87</td>
<td>96.2</td>
<td>79.3</td>
<td>0.86</td>
<td>90.7</td>
<td>70.9</td>
<td>0.61</td>
</tr>
<tr>
<td>ArcCHECK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>98.4</td>
<td>87.2</td>
<td>0.96</td>
<td>97.2</td>
<td>81.6</td>
<td>0.95</td>
<td>93.9</td>
<td>74.1</td>
<td>0.83</td>
</tr>
<tr>
<td>Delta4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>96.2</td>
<td>86.6</td>
<td>0.53</td>
<td>93.4</td>
<td>78.5</td>
<td>0.58</td>
<td>85.5</td>
<td>68.8</td>
<td>0.33</td>
</tr>
<tr>
<td>Gafchromic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>98.1</td>
<td>88.2</td>
<td>0.81</td>
<td>94.6</td>
<td>76.5</td>
<td>0.62</td>
<td>91.2</td>
<td>70.1</td>
<td>0.54</td>
</tr>
<tr>
<td>EPID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>97.7</td>
<td>77.4</td>
<td>0.82</td>
<td>96.2</td>
<td>66.3</td>
<td>0.84</td>
<td>93.6</td>
<td>59.1</td>
<td>0.82</td>
</tr>
</tbody>
</table>

The Delta4 measurement showed the lowest concordance correlation with the predicted gamma index. Given this difference, the Delta4 was re-analysed using 4%/3mm and 4%/4mm to see if there is better agreement to passing criteria of 3%/3mm in the other systems and the predicted evaluation.
6. Comparison of the gamma index in commercial systems

Figure 6.1  Cumulative histogram of gamma index analysis using passing criteria of 3%/3mm for (a) comparison of different software against the independent predicted calculation and (b) comparison of the measurements by the QA systems against the independent prediction. The software in the figure legend in (a) is listed in the same order as the associated hardware in (b).
6. Comparison of the gamma index in commercial systems

Figure 6.2  Scatter plot of gamma index analysis for each plan using passing criteria of 2%/2mm for (a) comparison of different software against the independent predicted calculation and (b) comparison of the experimental measurements by the QA systems against the independent prediction. The dashed line gives a 1:1 trend that would be expected for perfect agreement. The software in the figure legend in (a) is listed in the same order as the associated hardware in (b).
6.6.2 Calculated γ based on experimental linac measurements

All of the normal (unperturbed) plan measurements had a γ<1 passing rate of >98% using 3%/3mm in all systems. When using 2%/2mm, the passing rate for all the normal plans was greater than 93%. The 5mm systematic MLC positional modifications showed a visible strip of γ>1 for passing criteria 3%/3mm in all QA systems.

For the experimental measurements the trend against the predicted independent calculation is shown in the cumulative histogram in Figure 6.1b and the scatter plot in Figure 6.2b. For the measurements, a lower mean passing rate indicates greater response. Additionally, $\rho_c$ is given for the comparison between the measured plans and the independent predicted γ for each system in Table 6.1.

Figure 6.3 shows the cumulative histogram of the 4%/4mm, 4%/3mm, and 3%/3mm analysed in the Delta4 alongside the 3%/3mm results from the predicted gamma index calculation, and in Figure 6.4 the $\rho_c$ for the Delta4 vs the predicted is shown for the different passing criteria compared against the 3%/3mm $\rho_c$ for the other systems.

![Figure 6.3](image_url) Concordance correlation coefficients for the different systems for 3%/3mm. For the Delta4, the coefficient is also given for 4%/3mm and 4%/4mm.
Figure 6.4  Cumulative histogram of gamma index analysis for the Delta4 using passing criteria of 3%/3mm, 4%/3mm and 4%/4mm against the predicted 3%/3mm results.

Figure 6.5 shows a comparison between the gamma index distribution (using 3%/3mm) for the head & neck IMRT plan with a 5mm MLC positional modification measured by the 2D-Array, Delta4, ArcCHECK, EPID, and Gafchromic film and predicted by the independent gamma index calculation. For comparison separately, the predicted gamma index distribution in the SNC software vs the ArcCHECK measured gamma index map is shown in Figure 6.6.
6. Comparison of the gamma index in commercial systems

Figure 6.5 Gamma index distributions for the head & neck IMRT plan with 5mm MLC deliberate change. The passing criteria shown is 3%/3mm. Distributions are shown for the independent gamma index calculation, and the measured plans using the 2D-Array, Gafchromic film, EPID, ArcCHECK and Delta4.

The 2mm systematic MLC positional modifications were difficult to detect using 3%/3mm but were detectable for criteria of 2%/2mm. For the head & neck 2-arc RapidArc plan, none of the deviations were visible in the measurement and were also found to have a low impact on the expected gamma index maps. This is due to the plan having opposing collimator rotations on each arc to minimise the addition of inter-leaf leakage in the same spots, and therefore the deviations may have been largely masked. For the prostate & pelvic nodes RapidArc plan with collimator rotation changes,
measurements by all the QA systems at 3%/3mm gave a passing rate of >95% for deliberate 1 and 2 degree collimator rotation changes, and ranged from 77.0% - 88.2% in the presence of a 5 degree change.

Figure 6.6 Predicted vs measured gamma index points with $\gamma > 1$ for the ArcCHECK.

Figure 6.7 shows a comparison, for the measurements, of the percentage of detectors/pixels passing with a gamma index < 1, for passing criteria 2%/2mm, for the prostate IMRT plan with 2mm MLC positional modification, head & neck IMRT plan with 2mm MLC positional modification, prostate RapidArc with 2mm MLC positional modification, head & neck RapidArc with 2mm MLC positional modification and the prostate and pelvic nodes RapidArc plan with a collimator rotation error of 1 degree.
6. Comparison of the gamma index in commercial systems

Figure 6.7  Percentage of detectors/pixels passing with a gamma index < 1, for 2%/2mm, for the prostate IMRT plan with 2mm MLC positional modification, head & neck (H&N) IMRT plan with 2mm MLC positional modification, prostate RapidArc (RA) with 2mm MLC positional modification, H&N RA with 2mm MLC positional modification and the prostate and pelvic nodes (PPN) RA plan with a collimator (col.) rotation error of 1 degree.

6.7 DISCUSSION

This study has shown that various commercial software agree well with each other in calculating the predicted gamma index passing rates when delivery fluctuations and under-sampling effects due to the sparse detector resolution are removed, even at tight passing criteria of 2%/2mm (as shown in Table 6.1). Figure 6.1a illustrates the agreement for a passing rate of 3%/3mm using the cumulative histogram. The scatter plot in Figure 6.2a is given for tighter passing criteria of 2%/2mm and shows the trend against the expected 1:1 trend for perfect agreement. The trend at 2%/2mm shows that the software calculation tends to slightly over-estimate the passing rate with respect to the independent calculation, with the exception of one point where OmniPro predicted a lower passing rate for the prostate and pelvic node RapidArc plan with 5° collimator angle rotation change. Overall, the agreement is encouraging given the variability in detector array design from planar (2D-Array, EPID), helical (ArcCHECK) and cross-plane (Delta4) configurations as well as potential variation in implementing the gamma index calculation in the various software. There was also a good
agreement ($\rho_c > 0.92$ in all cases) against the independent 2D gamma index calculation. This agreement in the virtual calculation may be explained by the fact that none of these systems perform a true 3D gamma index calculation. Instead, generally the 3D dose cubes are re-sampled into 2D planes to compare against the measurements which are effectively 2D data. For example, in the ArcCHECK the measurement by the helical diode array is unrolled into a 2D plane and likewise the same is done with the predicted dose. In Verisoft there is a 3D option, but this takes into account the predicted dose planes above and below the measured 2D dose plane.

For the measured data it can be seen in Table 6.1 that the agreement against the predicted $\gamma$ reduces with tightening passing criteria and the variability between the different systems increases. This indicates that detector configuration and resolution have greater impact on the experimental calculation of $\gamma$. This is also seen in Figure 6.2b where there is more spread in the data compared to the graph for the predicted software analysis. This can be caused either by under-sampling of the dose distribution, blurring effects, or noise, or a combination of all. Moreover, in the theoretical approach, the high resolution combined with the comparison of two distributions with the same resolution mean that interpolations are not required. In particular the cases with fine resolution mean that for the DTA criteria used, the ratio between the grid spacing and the DTA will be less than the recommended one third [45]. For a real measurement using a sparse detector array, interpolation to a finer resolution will be required and/or re-sampling of the 3D dose distribution to the configuration of the detector array (as mentioned above) is required. These processes can introduce uncertainties particularly for a complex array configuration such as the Delta4. Practically, the Delta4 was found to have the lowest concordance coefficient based on measurements relative to the other systems for the same passing criteria, indicating lower agreement with the predicted gamma index as shown in Figure 6.1b and Figure 6.2b, and Table 6.1. As shown in Figure 6.3 and in Figure 6.4, it is possible to modify the passing criteria to get better agreement with the other systems, e.g. choosing 4%/3mm for the Delta4 was in better agreement with 3%/3mm in the other systems. However it should be noted that the lower agreement may indicate the possibility of potential false failures. Lower passing rates indicate that a system has greater measurement response to delivery deviations compared to the other systems and is therefore more likely to cause an investigation of failures in clinical practice. It is interesting to note that despite the high resolution of EPID and Gafchromic film being closest to the virtual measurements; their concordance correlation coefficients indicate weak correlation with the predicted $\gamma$. For the film, this is most likely due to intrinsic film heterogeneity causing noise artefacts combined with variability in scanning procedures, which can be enough to disrupt the gamma index analysis passing rate [91,124]. The decreasing concordance coefficient with tightening criteria supports the effect of noise in the film.
The EPID is also susceptible to small anomalies as reported by Gordon et al [131]. For EPID, the concordance coefficient appears to be stable across the different passing criteria suggesting the effects are systematic which may be correctable in future releases of the software and therefore improve the agreement [131]. Out of all the systems, the ArcCHECK measurements exhibited the closest statistical agreement with the predicted gamma index where for passing criteria of 3%/3mm and 3%/2mm $\rho_c$ was >0.95. The deliberately inserted modifications were designed to test the QA systems, however, the likelihood of this type of error occurring in clinical practice should be considered. An MLC motor is affected due to wear-and-tear, leading to a leaf travelling slower than expected and therefore the leaf lag behind the other leaves, in a way that would be similar to the deviations simulated in this study. The tolerance on the MLC control software (commonly 2mm on a Varian linac), means that there are generally two possible feedback scenarios: if possible, all the other leaves are slowed down and the dose rate is decreased to compensate for the slower leaf; or an interlock may be activated. Software errors may also lead to a mis-translation of the MLC positions. This study has demonstrated that in the event of subtle changes, as shown in Figure 6.5, Figure 6.6 and Figure 6.7, the detector array systems are able to detect some of these deviations if suitable passing criteria, such as 2%/2mm, are used. However, even lower passing criteria may be required for film, EPID, and ArcCHECK. Heilemann et al also concluded that 2%/2mm is necessary to detect positional deviations in MLCs for RapidArc deliveries [76].

This study has focussed on whether the measured gamma index calculation in various systems correlates with the predicted gamma index calculation. However, all systems have strengths and weaknesses and other aspects of these systems including usability, time taken to perform QA, ease of use and transportability, and so forth, were outside the scope of the present work. It should also be considered that all the systems were used according to manufacturer’s recommendations, and that in practice one may refine the methodology for using each to achieve optimal use of the system.

It is important to understand the sensitivity and limitations of the gamma index analysis combined with the equipment in use. For the same passing criteria, different devices and software combinations exhibit varying levels of agreement with the predicted analysis. When looking at the literature, it is important to note the type of equipment used and gamma index criteria. As shown in this study, passing criteria of 3%/3mm may not give the same results for measurements by different QA systems.
The Impact of the Gamma Index Calculation in 2D and 3D on Detector Array Resolution

7.1 Introduction

As previously discussed, detector arrays have been limited by their spatial resolution with most commercial systems limited to no better than 0.5cm centre-to-centre spacing. In chapters 4 and 6 the detector array systems have been relatively compared against Gafchromic film. Recently, a high resolution commercial system has been made available which has a centre-to-centre spacing of 0.25cm. This system is the PTW OCTAVIUS 1000 SRS detector array.

Furthermore, traditionally the γ index has been used to compare a 2D measured plane against a 3D dose distribution. There is a growing trend in using 3D back-projected EPID based in-vivo systems and independent calculations using Monte Carlo [132] to perform 3D versus 3D gamma index comparison. DVH-based analysis techniques have been proposed, however the gamma index continues to be prevalent. There have been quasi-3D commercial systems available such as the ArcCHECK and Delta4; however these have not constructed a true 3-dimensional dose distribution. Recently, software algorithm and hardware improvements have led to the possibility of using measured 2D data from commercial detector arrays to reconstruct a 3D-dose distribution and perform a volumetric comparison against the TPS. A limitation is that detector arrays have so far been limited by their spatial resolution which may affect the accuracy of the reconstructed 3D volume and subsequently the γ calculation due to under-sampling [126]. One commercial system that can use 2D measured data to reconstruct a 3D dose cube is the PTW OCTAVIUS® 4D [64,65]. At the time of writing, this can be used either with the OCTAVIUS® Detector 729 or with the high resolution OCTAVIUS® Detector 1000 SRS.

The purpose of this study was to extend the work from chapter 5 by comparing the PTW 2D-Array 729 directly against the OCTAVIUS 1000 SRS to check the impact of the gamma index passing rate on
two detector array systems with significant difference in resolution. An additional aim was to assess whether the gamma index passing rate in a 2D plane measurement was a suitable surrogate for a 3D volumetric gamma index.

7.2 METHOD

7.2.1 Equipment

Due to its design, it was not possible to use the OCTAVIUS 1000SRS detector array within the octagonal OCTAVIUS II phantom to measure a composite 2D plane to directly compare against previous measurements using the 2D-ARRAY 729. Instead, the 1000SRS array required the use of the PTW OCTAVIUS 4D phantom. Measurements were also performed using the 729 array within the OCTAVIUS 4D.

7.2.1.1 Cross-calibration procedure

At the time of measurements on the linac, the detector arrays were cross-calibrated in the OCTAVIUS 4D phantom; in this procedure, a known dose was delivered and the response of the central detector was used to calculate a cross-calibration factor. This factor was applied to the entire matrix of detectors [94].

7.3 DELIBERATE PLAN MODIFICATION TESTS

The deliberate change tests for the prostate IMRT and RapidArc plans used in Chapter 4 and 6 were used. Following the same methodology used previously (Chapter 6), the plans were measured with the OCTAVIUS4D with the SRS1000 array and 729 array in situ. The 2D coronal plane measurements that were previously performed using the 2D-ARRAY 729 within the OCTAVIUS II phantom were used.

The 1000SRS array is limited to measuring over an active area on 11cm x 11cm. Therefore the 3D dose cylinder is limited to 11cm length and 5.5cm radius. In order to maintain a consistent analysis, the 3D dose reconstruction was performed in a cylinder of 11cm length and 5.5cm radius for the 2D-ARRAY 729 and SRS1000. It was possible to manually specify the dimension of the reconstructed dose distribution within the Verisoft software, as shown in Figure 7.1. For the 2D coronal plane measurements using the 2D-ARRAY 729 in the OCTAVIUS II phantom, the region of interest for the analysis was set to 11cm x 11cm.
The expected 2D and 3D γ pass rates were simulated by exporting the normal plan and perturbed plan predicted dose distributions in DICOM format. The predicted γ was calculated, using the bespoke Matlab software, for a 2D coronal-plane as well as in 3D. Perturbed plan measurements (2D coronal-plane and 3D-reconstructed-dose) were evaluated against the normal dose distribution. The same cylinder and plane dimensions, above, were applied to the simulated 3D and 2D dose distributions respectively.

For evaluations, the global γ was used with passing criteria of 3%/2mm, and a 20% threshold. The γ results based on measurements were compared against the 3D volume predicted analysis. The concordance correlation coefficient, $\rho_c$, was used to assess agreement between the analyses.

![Image of parameter specification window](image)

Figure 7.1  The parameter specification window for 3D dose reconstruction in Verisoft V6.0.

### 7.4 Results
A summary of mean percentage of points passing with γ<1 is given in Table 7.1. There was good agreement between the predicted γ for the 2D coronal plane and 3D volume analysis ($\rho_c$>0.90). The
measured 2D and 3D data had a lower agreement with the predicted and had consistently lower pass-rates, see Figure 7.2. However, there was statistically strong agreement between the 3D 1000SRS measured pass-rate and 3D predicted pass-rate ($\rho_c=0.93$). The OCTAVIUS 729 used in the OCTAVIUS 4D phantom to reconstruct a 3D dose distribution had the worst agreement of $\rho_c = 0.28$.

Table 7.1 Summary of mean and minimum gamma index passing criteria for each system for the range of changes introduced. The concordance correlation coefficient, $\rho_c$, is also given assessing agreement with the predicted 3D volume gamma index pass-rate.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>min</th>
<th>$\rho_c$</th>
<th>Pearson $\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Predicted</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D volume</td>
<td>98.8</td>
<td>92.4</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2D coronal plane</td>
<td>98.6</td>
<td>91.8</td>
<td>0.98</td>
<td>0.99</td>
</tr>
<tr>
<td><strong>Measured</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OCTAVIUS729 2D corona</td>
<td>97.8</td>
<td>86.9</td>
<td>0.81</td>
<td>0.99</td>
</tr>
<tr>
<td>OCTAVIUS 1000 SRS 3D volume</td>
<td>98.1</td>
<td>91.5</td>
<td>0.93</td>
<td>0.98</td>
</tr>
<tr>
<td>OCTAVIUS729 3D volume</td>
<td>97.2</td>
<td>94.6</td>
<td>0.34</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Figure 7.3 shows a comparison of 2D dose profiles (plotted as discrete points) for the RapidArc prostate normal plan measurement from the 3D reconstructed dose distributions using the 729 and 1000SRS array. The profile from the TPS predicted dose distribution is plotted for comparison.
7. Impact of 2D and 3D gamma index on detector array resolution

Figure 7.2  Global gamma index passing rates for 3%/2mm for the different equipment and simulated analysis.

Figure 7.3  2D dose profile for the 3D reconstructed measured dose from the 729 and 1000\textsuperscript{SRS} arrays compared against the TPS dose profile.
7. Impact of 2D and 3D gamma index on detector array resolution

Figure 7.4 2D coronal planes. **Top row:** TPS, SRS1000 in OCTAVIUS 4D, OCTAVIUS 729 in OCTAVIUS 4D, and 2D-ARRAY 729 in OCTAVIUS II dose distribution for the plan with 5mm MLC positional modification. **Bottom row:** 3%/3mm γ distribution for the plan with 5mm MLC positional modification in Matlab predicted, OCTAVIUS-1000SRS, OCTAVIUS 729, and coronal 2D-ARRAY.

7.5 DISCUSSION

This study has shown that the impact of the 2D and 3D gamma index can be significant depending on the spacing of detectors within an array and the way that it is used. In the case of a planar measurement using the sparsely arranged 2D-ARRAY 729 (as originally intended for this piece of equipment) the gamma index passing rate was in reasonable agreement with the predicted 3D volumetric passing rate. However when the 2D-ARRAY 729 was used in the OCTAVIUS 4D phantom to reconstruct a 3D dose distribution it gave the worst overall agreement at $\rho_r = 0.28$. The gamma index distribution in Figure 7.4 indicates that this is likely due to artificial bands of failed points. The reason for these is given in Figure 7.3 where it is clear that there has been incorrect interpolation at the profile edges due to the low resolution which can result in the interpolated dose points in those regions appearing to have an artificially lower dose than the TPS dose. This problem is shown to disappear when using the higher resolution 1000SRS detector array where the dose reconstruction is more accurate, leading to a statistically stronger agreement.

Interestingly the agreement for the passing rate using the 2D-ARRAY 729 within the OCTAVIUS4D phantom improves to $\rho_r = 0.80$ when comparing the measured analysis using 3%/3mm against the
predicted 3D analysis using 3%/2mm. This is similar to the trend found with the Delta4 in Chapter 5, where the Delta4 appeared to have greater sensitivity to measurements, but with the lowest agreement with the predicted gamma index. Therefore similar interpolation issues, as found in this study, may be apparent in the Delta4’s proprietary gamma index calculation. However the ‘black box’ natures of the Delta4 calculation, means it is not possible to investigate further.

The development of the 1000SRS is a promising step in the detector array technology. With a spacing of 2.5mm in the central 5.5cm x 5.5cm area, this is similar to the typical grid spacing in treatment planning systems providing better accuracy in the dose reconstruction as demonstrated in this study. The major limitation is the size of the array which would be limited to small treatment regions. This study shows that sparse detector arrays may increase measurement uncertainty when used to reconstruct 3D dose distributions, however further studies with a number of different plans with different geometries are required for confirmation.
8

DEVELOPMENT OF A METHODOLOGY FOR USING A DETECTOR ARRAY IN A NATIONAL ROTATIONAL IMRT DOSIMETRY AUDIT

8.1 BACKGROUND AND RATIONALE

Radiotherapy dosimetry audits allow for the testing of procedures and the identification of errors [31–33]. Dosimetry audits range in complexity from measuring machine output under reference conditions [31,32,34–38] to complex radiotherapy such as intensity modulated radiotherapy (IMRT) measurements [39–43]. As the complexity of treatment technology continues to increase so do the potential uncertainties and inaccurate dose delivery can have clinical implications [29,30]. Furthermore, there is an important role of dosimetry audits within clinical trials where there is some evidence that poor dosimetry may affect clinical outcomes [133–135]. Dosimetry audits can be classified into three levels as described by [136]; these are:

**Level I:** independent measurement of linear accelerator output under reference conditions in a regular phantom (for example solid water blocks)

**Level II:** independent dose measurement with variations in the irradiation configuration (e.g. IMRT) and position of measurement in a regular phantom

**Level III:** independent measurement of dose in an anthropomorphic phantom that is planned and treated as similar to a patient as possible.

These can be further broken down into two categories:

1. **Site visit:** taking equipment to a neighbouring centre or a range of centres and performing an independent measurement.

2. **Postal audit:** a centrally organised audit using dosimeters such as TLDs or alanine, for example the IAEA/WHO system [36]
In the UK, Level I category 1 dosimetry audits are organised by the Institute of Physics and Engineering in Medicine (IPEM) through 8 regional audit subgroups, and the National Physical Laboratory (NPL). National dosimetry audits for Mega-voltage (MV) beams and electrons have previously been performed [31,32] and set the benchmark for Level 1 audits in the UK.

Postal audits are usually favoured for international audits such as those organised by the International Atomic Energy Agency [IAEA] [36] and the European Society for Therapeutic Radiation Oncology (ESTRO) [35].

Since the introduction of commercial volumetric modulated arc radiotherapy (VMAT) in 2008 [9,26] there has been a fast uptake of this complex technology, as well as an increase in the number of centres with Helical Tomotherapy. A survey of United Kingdom (UK) cancer centres in July 2010 indicated that around 30% were treating with some form of rotational radiotherapy (Varian RapidArc [Varian Medical Systems Inc., Palo Alto, CA], Elekta VMAT [Elekta AB, Stockholm, Sweden] or Helical Tomotherapy [Accuray-Tomotherapy, Madison, WI]), and that this would increase to 50% by the end of 2011 [137]. Conventional methods such as individual ionization chamber point dose measurements and film dosimetry are time consuming. Therefore, a novel methodology was required in order to perform a large-scale dosimetry audit of this emerging technology.

In this study, development of a methodology for using a commercial detector array in a dosimetry audit of complex rotational radiotherapy has been undertaken. This work was carried out in collaboration with the National Physical Laboratory (NPL), NCRI Radiotherapy Trials QA group and the Institute of Physics and Engineering in Medicine. The methodology has been developed by evaluating the approach in ten cancer centres treating with some form of rotational radiotherapy. Comparison was made between measurements using the detector array and those performed using 0.125 cm$^3$ ionization chambers, alanine and EBT2 Gafchromic film (International Specialty Products, Wayne NJ).

### 8.2 Methods and Materials

Ten UK cancer centres that had already begun treating with a form of rotational radiotherapy (Varian RapidArc, Elekta VMAT or Helical Tomotherapy) were visited in the period between June 2011 and November 2011. Of those centres, five used Varian RapidArc, two centres used Pinnacle SmartArc (Philips Healthcare, Eindhoven, Netherlands) delivered on Elekta and three had
Helical Tomotherapy. One centre also chose to use Pinnacle SmartArc, in addition to RapidArc, delivered on a Varian linear accelerator. A full summary is given in Table 8.1.

<table>
<thead>
<tr>
<th>Centre</th>
<th>Machine</th>
<th>Modality</th>
<th>Delivery type</th>
<th>TPS and version</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Varian Clinac 2100C/D</td>
<td>RapidArc</td>
<td>2-Arc</td>
<td>Varian Eclipse v8.6.15</td>
</tr>
<tr>
<td>2</td>
<td>Varian Clinac 2100C/D</td>
<td>RapidArc</td>
<td>2-Arc</td>
<td>Varian Eclipse v8.6.17</td>
</tr>
<tr>
<td>3</td>
<td>Elekta Synergy</td>
<td>SmartArc (VMAT)</td>
<td>2-Arc</td>
<td>Philips Pinnacle3 v9.0</td>
</tr>
<tr>
<td>4</td>
<td>Varian Clinac 2300iX</td>
<td>RapidArc</td>
<td>2-Arc</td>
<td>Varian Eclipse v10.0.34</td>
</tr>
<tr>
<td>5</td>
<td>Varian Clinac 2300iX</td>
<td>SmartArc (VMAT)</td>
<td>2-Arc</td>
<td>Philips Pinnacle3 v9.0</td>
</tr>
<tr>
<td>6</td>
<td>Elekta Synergy</td>
<td>SmartArc (VMAT)</td>
<td>Single-Arc</td>
<td>Philips Pinnacle3 v9.0</td>
</tr>
<tr>
<td>7</td>
<td>Tomotherapy HiArt</td>
<td>Tomotherapy</td>
<td>Helical</td>
<td>HiArt v4.0</td>
</tr>
<tr>
<td>8</td>
<td>Tomotherapy HiArt</td>
<td>Tomotherapy</td>
<td>Helical</td>
<td>HiArt v3.0</td>
</tr>
<tr>
<td>9</td>
<td>Varian Clinac 2300iX</td>
<td>RapidArc</td>
<td>2-Arc</td>
<td>Varian Eclipse v8.9.15</td>
</tr>
<tr>
<td>10</td>
<td>Tomotherapy HiArt</td>
<td>Tomotherapy</td>
<td>Helical</td>
<td>HiArt v4.0</td>
</tr>
</tbody>
</table>

8.2.1 The RTTQA 3DTPS Test for RapidArc, VMAT and Tomotherapy

All centres taking part were asked to plan a three-dimensional (3D) treatment planning system (TPS) test, that had been designed by the UK National Cancer Research Institute (NCRI) Radiotherapy Trials Quality Assurance (RTTQA) group specifically for the rotational radiotherapy delivery techniques, henceforth referred to as the 3DTPS Test [138]. This test is based on the multi-target benchmarking test reported in AAPM report 119 [139] and the RTTQA clinical trial credentialing programme first described by Clark et al [41,140]. Development of this test and its application has previously been described [138]. In summary, this test checks multi-leaf collimator (MLC) positioning, transmission and relative dose levels in transverse, coronal and sagittal planes. The 3DTPS Test, shown in Figure 8.1, is comprised of a virtual 20 cm diameter cylindrical homogeneous water phantom and a set of clinically relevant volumes. These volumes include five planning target volumes (PTVs) and a single organ at risk (OAR). PTV2 is the primary target volume and is prescribed a dose of 25 Gy in 10 fractions. PTV1 and PTV4 are prescribed 15 Gy, and PTV3 and PTV5 prescribed 20 Gy. The OAR was required to be kept below 10 Gy [138].
Centres were asked to use their common clinical parameters to plan the 3DTPS Test. Each plan was required to be optimised following the guidelines in the International Commission on Radiation Units and Measurements (ICRU) report 83 [112] to achieve the dose constraints for each of the PTVs whilst minimising dose to the OAR and remaining volume as far as possible [138] and were assessed by the RTTQA group.

![Image](image_url)

**Figure 8.1** Transverse slice image of the 3DTPS test (left). The dashed lines indicate the planes measured. The black dots indicate dose points, note that two points were taken in PTV2 (one in the coronal and one in the sagittal plane). The image on the right shows a 3D reconstruction of the test with all PTVs and OAR labelled.

<table>
<thead>
<tr>
<th>VOLUME</th>
<th>Mean Dose</th>
<th>Min Dose</th>
<th>Max Dose</th>
</tr>
</thead>
<tbody>
<tr>
<td>OAR</td>
<td>N/A</td>
<td>N/A</td>
<td>$D_{\text{max}} &lt; 10\text{Gy}$</td>
</tr>
<tr>
<td>PTV1</td>
<td>15Gy ±0.5Gy</td>
<td>$D_{99%} &gt; 13.5\text{Gy}$</td>
<td>$D_{10%} &lt; 16.5\text{Gy}$</td>
</tr>
<tr>
<td>PTV2 (Primary PTV)</td>
<td>25Gy ±0.5Gy</td>
<td>$D_{99%} &gt; 22.5\text{Gy}$</td>
<td>$D_{1%} &lt; 26.75\text{Gy}$</td>
</tr>
<tr>
<td>PTV3</td>
<td>20Gy ±0.5Gy</td>
<td>$D_{99%} &gt; 18\text{Gy}$</td>
<td>$D_{10%} &lt; 22\text{Gy}$</td>
</tr>
<tr>
<td>PTV4</td>
<td>15Gy ±1Gy</td>
<td>$D_{99%} &gt; 13.5\text{Gy}$</td>
<td>$D_{10%} &lt; 18\text{Gy}$</td>
</tr>
<tr>
<td>PTV5</td>
<td>20Gy ±0.5Gy</td>
<td>$D_{99%} &gt; 18\text{Gy}$</td>
<td>$D_{10%} &lt; 22\text{Gy}$</td>
</tr>
</tbody>
</table>

$D_{\text{max}}$ = Maximum dose, $D_{99\%}$ = Dose received by 99% of the volume, $D_{10\%}$ = Dose received by 10% of the volume, $D_{1\%}$ = Dose received by 1% of the volume.

### 8.2.2 Choice of detector array and phantom combination

The choice of the detector array system for this study depended on the following criteria:
1. Designed to be able to measure rotational radiotherapy

2. Be robust and transportable

3. Allow direct comparison with ionization chambers, film and alanine

The PTW OCTAVIUS II phantom and 2D-Array seven29 combination fulfilled all of the criteria identified above and was also a system that members of the audit team were already familiar with in their respective clinical institutions [94]. Figure 8.2 shows the versatility of the OCTAVIUS II system. Two inserts from the manufacturer were available, one of which was designed to allow a PTW semiflex 0.125 cm³ ionization chamber to be used in nine different positions. The design was such that the measurement depth was the same as that of the detectors within the 2D-Array and ensured that the position of the collecting volume of the semiflex chambers corresponded with the nine central detectors in the central row of the 2D-Array. Another insert held 20 cm x 25.4 cm Gafchromic film sheets at the same measuring plane as the 2D-Array. An in-house insert was made for the alanine pellets and was designed to provide identical measuring positions as available for the semiflex chamber insert.

Figure 8.2 The OCTAVIUS II system with (a) 2D-ARRAY 729, (b) Gafchromic EBT2 film insert, (c) PTW semiflex 0.125cc ion chamber insert, and (d) alanine insert. Note that for the 2D-ARRAY 729, the base of the OCTAVIUS phantom has a compensation gap to correct for the inherent under-response of the array; this is indicated by the black
smile on the outside. For the other dosimetry systems, a solid base is used (indicated by the red smile).

8.2.3 Semiflex ionization chambers

Two PTW Semiflex 0.125 cm$^3$ ionization chambers were taken on each audit visit. The ionization chambers were calibrated in terms of absolute dose to water traceable to the primary standard held at the National Physical Laboratory (NPL, Teddington, UK) for a range of quality indices (QI; TPR20/10). The host centres provided information regarding the QI prior to each visit. Absorbed dose to water calibration factors for the given QI were extracted from the NPL calibration curves. Temperature and pressure were corrected using a calibrated thermometer and barometer from the NPL.

8.2.4 Gafchromic EBT2 film

Gafchromic EBT2 20 cm x 25.4 cm film sheets were used. The orientation of the film was always consistent for all measurements; each film had a mark at manufacture to allow consistency in setup. All Gafchromic films were processed and analysed by another member of the audit team (Yat Tsang) at Mount Vernon Hospital at least 24 hours after exposure. Films were scanned using the Epson Espression 10000 XL flatbed (Seiko Epson Corp., Nagano, Japan) colour scanner at a resolution of 75 dpi, using the red channel [91]. The films were all scanned in the same orientation and a jig was used to place the films in the same part of the scanner to minimize any displacement effects and to use the optimum part of the scanner which was determined through the commissioning process. A uniformity correction was applied by scanning a blank film from each batch. A calibration curve for each Gafchromic film batch was determined for a range of doses between 0 and 600 cGy at Mount Vernon Hospital.

8.3 ALANINE PELLETS

Alanine pellets were provided by the NPL. These were chosen for three reasons:

1. Absolute dosimetry, direct traceability to the primary graphite calorimeter held at NPL
2. Energy independence
3. To allow the possibility of a postal audit if necessary

Each pellet was 2.4 mm in thickness and 4.5 mm in diameter. The alanine batch had calibration factors determined at the NPL. The pellets used required a dose of around 1000 cGy to reduce the
uncertainty associated with the measurement. At this dose level the uncertainty in the measurement was 0.85% (coverage factor, k=1). The pellet-to-pellet reproducibility was 0.5% (k=1) [43,141]. In all measurement cases, three alanine pellets were stacked together within each location in the jig. All pellets were processed by the NPL using electron paramagnetic resonance spectrometry (EPR) [141,142].

8.3.1 Miscellaneous equipment

Barometers and thermometers, calibrated at the NPL, were taken on each audit visit.

8.4 SEMIFLEX IONIZATION CHAMBER OUTPUT CALIBRATION

On conventional linear accelerators, standard output measurements were carried out following the UK Institute of Physics and Engineering in Medicine (IPEM) code of practice [143] by setting up the ionization chamber in solid water and delivering 200 MU to 5 cm depth at 95 cm or 100 cm focus-to-surface-distance (FSD), according to the host centre protocol, for a 10x10 cm field size.

The Tomotherapy centres used a helical plan for their absolute dose calibration, which delivered a homogeneous dose of 200 cGy to a 6 cm diameter by 6 cm length cylindrical PTV in the Tomotherapy Virtual Water™ phantom (commonly referred to as the ‘cheese’ phantom) [144]. To perform an independent output, users were asked to generate a delivery QA (DQA) plan in the OCTAVIUS scan using the cheese phantom helical output plan. The expected dose to the central semiflex ionization chamber position was determined. The output difference was thus determined by comparing the measured semiflex with the expected dose.

In all cases, the average of at least 5 measurements was calculated. The output was measured at the start and end of each audit, and all ionization chamber measurements were subsequently corrected for the average output. The output was also determined by each host centre using their own ionization chamber and phantom to cross-reference with the output measured by the audit team. The difference between the outputs was expected to be within ±1%.

8.5 2D-ARRAY CROSS-CALIBRATION

The 2D-Array was cross-calibrated in the OCTAVIUS phantom. The water equivalent depth to the centre of the chambers in the array was 16 cm. An isocentric setup was used. On the conventional linear accelerators a 10x10 cm single static field was used. The host centre was asked to provide the
monitor units required to deliver 200 cGy to the chambers in this setup condition. For Tomotherapy, this was done by delivering a helical output plan in the same way as the ionization chamber measurement. Prior to cross calibrating the 2D-Array, the semiflex ionization chamber and alanine pellets were inserted into the OCTAVIUS phantom to independently measure the expected dose in this condition and minimise uncertainty in the array measurement.

### 8.6 Verification Plan Creation

Each centre was provided with four sets of computed tomography (CT) scans of the OCTAVIUS phantom. These were as follows:

1. OCTAVIUS phantom with 2D-Array inserted, with the detector plane setup in the coronal (horizontal) orientation

2. OCTAVIUS phantom with 2D-Array inserted, with detector plane setup in the sagittal (vertical) orientation

3. OCTAVIUS phantom with homogeneous insert for ionization chamber, film, and alanine dosimetry, with the detector plane setup in the coronal (horizontal) orientation

4. OCTAVIUS phantom with homogeneous insert for ionization chamber, film, and alanine dosimetry, with the detector plane setup in the sagittal (vertical) orientation

Centres were provided with CT number to relative electron density and mass density calibration curves and were instructed to import the appropriate curve into their TPS where appropriate. This was not a mandatory step as the uncertainty was estimated to be within 0.5%. Each centre was instructed to apply their normal procedure for correcting for the couch; e.g. inserting a couch structure in the planning system. The collecting volumes of the ionization chambers (0.125 cm³) in the nine different possible positions within the insert were contoured onto the coronal and sagittal CT scans with the homogeneous insert to enable calculation of the predicted dose for the semiflex chamber and alanine.

Centres were given detailed instructions to ensure that the position of the 3DTPS Test dose distribution relative to the phantom was consistently reproduced, and thus the dose planes were measured in the same part of the plan from centre-to-centre. The dose planes that were measured to sample the regions of interest using all dosimetric methods are shown in Figure 8.3. There were two coronal (horizontal) planes and a sagittal (vertical) plane. The first coronal plane directly
intersected PTV1, PTV2, PTV4 and PTV5. The second coronal plane was 4 cm posterior with respect to the first, and intersected the OAR, PTV1 and PTV3. The sagittal plane intersected PTV2, PTV4, PTV5 and the OAR.

Six separate point dose locations were chosen, to sample different dose levels in the 3DTPS test. In the first coronal plane, these were a central point in PTV2 and a point within PTV1. In the sagittal orientation, a point was recorded in PTV2 and in the OAR. Centres were instructed to use the dose-volume-histogram (DVH) in their TPS to record the absolute mean dose and standard deviation for one fraction for each chamber. For the alanine measurements, at least 9 fraction deliveries were required to deliver at least 1000 cGy to the pellets. The average dose recorded by the pellets was therefore rescaled to a single fraction delivery to allow a direct comparison with the semiflex and 2D-Array.

All verification plans were returned for independent evaluation using the Visualization and Organization of Data for Cancer Analysis (VODCA) independent evaluation software version 4.3.3 [145]. DICOM dose cubes were exported from each plan to be used for analysis. Predicted doses for the ionization chamber and alanine were calculated using the TPS and independently verified using VODCA.

8.6.1 Measurement planes

Each separate OCTAVIUS scan had the 3DTPS volumes registered on to it relative to the 3DTPS phantom. This allowed for consistency in setting up the isocentre position so that the dose distribution is reconstructed in the same position on the OCTAVIUS phantom across different centres. The dose planes that were measured to sample the regions of interest using all dosimetric methods are shown in Figure 8.3. There were two coronal planes. The first coronal plane directly intersected PTV1, PTV2, PTV4 and PTV5. The second coronal plane was 4cm posterior with respect to the first, and intersected the OAR, PTV1 and PTV3. There was a sagittal plane which intersected PTV2, PTV4, PTV5 and the OAR.
Centres were given detailed instructions regarding the generation of verification plans to ensure that the dose distribution relative to the 3DTPS volumes on the OCTAVIUS phantom was consistent from centre-to-centre. In total six separate verification plans were generated for the 3DTPS test, Figure 8.4 shows how the dose distributions should look for the main planes to be measured. In total six separate verification plans were generated for the 3DTPS test:

1. Coronal plan with 2D Array inserted
2. Coronal plan with 2D Array inserted, isocentre 4cm ANT
3. Sagittal plan with 2D Array inserted
4. Coronal plan with homogeneous insert
5. Coronal plan with homogeneous insert, isocentre 4cm ANT
6. Sagittal plan with homogeneous insert
8.6.2 Individual dose point measurement

For the predicted dose for semiflex ionization chamber and alanine, the CT scans with the homogeneous insert had contoured on them the collecting volume of the PTW Semiflex ionization chambers. In both the sagittal and coronal CT scans, the chamber positions were contoured and labelled ‘Chamber 1 – 9’. Centres were instructed to use the DVH in their TPS to record the absolute mean dose and standard deviation for 1 fraction for each relevant chamber for the coronal and sagittal plans.

A total of 6 separate semiflex ionization chamber points were chosen, each in a different position to sample different dose levels in the 3DTPS test. In the 1st coronal plane, two points were recorded, a central point in the main PTV (PTV2) and a point within PTV1. In the 2nd coronal plane, a point was recorded in PTV1 and PTV3. In the sagittal orientation, a point was recorded in the main PTV and in the OAR. All are shown visually in Figure 8.5. Alanine measurements were also carried out in the same points. In addition, the individual chambers in the 2D-ARRAY 729 corresponding to the same
position as the semiflex ionization chambers were recorded to perform a direct comparison with the ionization chambers and alanine.

Figure 8.5  Point dose measurements in (a) coronal 1, (b) coronal 2, and (c) sagittal plane. Alanine measurements were carried out in (a) and (c).

8.7 ANALYSIS OF DATA

To evaluate the 2D-Array against the other systems of dosimetry, data analysis was carried out as described below.

8.7.1 Comparison of the 2D-Array against semiflex ion chamber and alanine

As there was a consistent approach for creating the 3DTPS verification plans, it was possible to make a direct comparison between the semiflex ionization chamber and alanine measured dose with the dose recorded by the corresponding individual detectors within the 2D-Array. The statistical agreement between the measurements by the three systems was tested by calculating the concordance correlation coefficient, $\rho_c$ [123], and generating Bland-Altman plots [146]. The Bland-Altman plot is achieved by plotting the absolute dose difference between two systems as a function of the mean dose measurement of the two systems. The 95% levels of agreement (i.e. the mean difference ± 1.96 s.d.) are also plotted. The R statistical software package was used [147] (R Core Team, University of Auckland, NZ). Furthermore, the measured dose points for the alanine were compared against the corresponding predicted dose from the TPS to provide a further comparison between the different systems.

8.7.2 Comparison between 2D-Array and Gafchromic film planes

To compare 2D-Array and Gafchromic film plane measurements, the global gamma index ($\gamma$) method of evaluation [45,48] was used with a 20% threshold relative to a point selected in a high dose, low
gradient region. The gamma index evaluation was performed taking into account the 3D dose distribution; i.e. the complete TPS dose cube, not only the dose distribution of the measured plane. Various criteria for $\gamma$ were analysed, including the commonly used 3% dose difference (DD) and 3mm distance-to-agreement (DTA) criteria. For the 2D-Array, analysis was performed using the PTW Verisoft software version 4.1. Absolute $\gamma$ analysis was performed for the 2D-Array; the dose point for the dose difference criteria was chosen in a high dose, low gradient region. In the absolute analysis neither the dose cube nor the 2D-Array were normalized. For Gafchromic film, the analysis was performed using the Scanditronix Wellhöfer OmniPro™ I’mRT software version 1.7 (IBA Dosimetry GmbH, Schwarzenbruch, Germany). Both the film and plan data were normalized to a point in a high-dose low-gradient region, to perform a relative comparison, as is currently used for dosimetry audit using film in the UK [41,43]. This procedure is commonly used for film analysis due to the known difficulty in performing an absolute dose calibration for film [68].

8.7.3 Reporting of data to host centres

For each audit visit, a report of the results was compiled. This included all output measurements, semiflex ionization chamber and alanine point dose results along with results from the corresponding ionization chamber within the 2D-ARRAY 729, and gamma index analysis for both the 2D-ARRAY 729 and Gafchromic film. Results for criteria of 4%/4mm, 3%/3mm and 2%/2mm were included, as well as gamma index distribution maps.

8.8 Timing study

During each visit, start and end times of events were recorded. Events recorded were categorised into: initial equipment setup, standard output measurements by audit team and host centre, output in OCTAVIUS phantom using semiflex chamber and alanine, 2D-ARRAY 729 cross-calibration, 2D-ARRAY 729 plan measurements (including online analysis), plan measurements using ionization chamber, Gafchromic film and alanine measurements for the 3DTPS test, output measurements at the end of the day. Miscellaneous events such as: set up between coronal and sagittal OCTAVIUS phantom orientation and time taken to swap the base of the OCTAVIUS after the 2D-ARRAY 729 measurements.

8.9 Results

In total 36 2D-Array and Gafchromic film planes were measured, as well as 72 ionization chamber points and 40 sets of alanine measurements. After processing the film, it was discovered that 9 out
of the 36 (25%) appeared to have significant artefacts that were traced back to a single batch. These films were subsequently discounted from further analysis. The difference between the standard ion chamber output measured by the audit team and host centre was within ±1% in all cases.

8.9.1 Comparison of individual 2D-Array chambers against semiflex ion chamber and alanine

The mean percentage dose difference between the 2D-Array and semiflex ionization chamber and alanine points was -1.6% ± 1.8% (mean ± one standard deviation) and -1.5% ± 1.8% respectively. It should be noted that this data includes points measured in a low dose high gradient region. However, removing the low dose high dose gradient data resulted in a mean difference of -1.1% ± 1.1% and -0.8% ± 1.1%, respectively for the 2D-Array against semiflex and alanine. The difference between the semiflex dose points and alanine was found to be narrower at 0.5% ± 0.7%. Figure 8.6 has scatter plots showing the trend between the 2D-Array doses against the semiflex ionization chambers and alanine doses, respectively, and between the semiflex and alanine doses. The corresponding Bland-Altman plots showing, along with 95% levels of agreement are also plotted. For the 2D-Array versus semiflex data, it can be seen that 95% of 2D-Array points measured were within -6.0 cGy and 2.4 cGy absolute dose differences. For the 2D-Array vs alanine, these were within -5.6 cGy and 2.3 cGy, and for ion chamber vs alanine this was narrower at -1.1 cGy and 2.3 cGy. There was a statistically significant concordance correlation coefficient for all the comparisons (ρc > 0.998, p<0.001 in all cases).

A histogram for the percentage difference between the measured dose and the predicted dose for each of the 2D-Array, semiflex ionization chamber, and alanine data is shown in Figure 8.7. It was found that the mean difference for the 2D-Array was +0.4% ± 3.0%, for the semiflex ionization chambers it was +0.2% ± 2.9%, and for alanine it was -0.4% ± 3.6%. In the high dose PTV region only (i.e. PTV2 in the 3DTPS test), this was -0.1% ± 1.7% for 2D-Array, -0.2% ± 1.8% for semiflex, and -0.6% ± 1.6% for alanine. It should be noted that since this data was normally distributed with random variations as seen in Figure 8.7 that it wasn’t possible to conduct a meaningful correlation analysis.
Figure 8.6  Scatter plots and Bland-Altman agreement plots for (a & b) individual 2D-ARRAY 729 detector dose against semiflex ion chamber measurements, (c & d) individual 2D-ARRAY 729 detector dose against alanine measurements, and (e & f) semiflex ion
chamber measurements against alanine measurements. On each scatter plot, perfect agreement is indicate by the dashed line. On each Bland-Altman plot, upper and lower 95% Limits of agreement are shown using dotted lines and mean difference using bold line.

8.9.2 Comparison between 2D-Array and Gafchromic film

A summary of the mean percentage of detectors/pixels with gamma index <1 for passing criteria for the 2D-Array and Gafchromic film planes for passing criteria 4%/4mm, 3%/3mm, and 2%/2mm is given in Table 8.3. The percentage of 2D-Array and Gafchromic film planes achieving ≥95% of detectors or pixels with a γ<1, P95%, is also given.

Table 8.3 Range of detectors/pixels passing with γ<1 for the 2D-Array and Gafchromic film planes for passing criteria 4%/4mm, 3%/3mm, and 2%/2mm. The percentage of 2D-Array and Gafchromic film planes achieving ≥95% of detectors or pixels with a γ<1, P95%, is also given.

<table>
<thead>
<tr>
<th>Device</th>
<th>4%/4mm</th>
<th>3%/3mm</th>
<th>2%/2mm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Range</td>
<td>P95%</td>
<td>Range</td>
</tr>
<tr>
<td>2D-Array</td>
<td>95.3 – 100</td>
<td>100</td>
<td>86.3 – 100</td>
</tr>
<tr>
<td>Gafchromic film</td>
<td>97.7 – 100</td>
<td>100</td>
<td>93.8 – 100</td>
</tr>
</tbody>
</table>

Figure 8.7 shows the histogram of the 2D-Array and Gafchromic film results for the commonly used passing criteria of 3%/3mm. When using the passing criteria of 3%/3mm, the P95% is 83.3% for 2D-Array and 88.9% for Gafchromic film. For this passing criteria, the plane that had the lowest detector pass-rate for the 2D-Array had 72.6% of the detectors with γ< 1, and the film plane with the lowest pixel pass-rate for Gafchromic film had 94.0% pixels with γ<1. The 2D-Array gave an absolute gamma index comparison whereas Gafchromic relied on a relative one, therefore the passing rate in the latter would be expected to be higher as dose differences are normalised out. Figure 8.8 shows an example of where there were 2D-Array failures for passing criteria of 3%/3mm due to a >3% dose deviation in the PTV region. The Gafchromic film passes as this dose deviation has been cancelled out in the relative gamma analysis.
Figure 8.7 Relative frequency histograms showing (a) percentage difference between measured dose and predicted dose for the 2D-Array, semiflex and alanine, and (b) 2D-ARRAY 729 versus Gafchromic film gamma index analysis calculated at 3%/3mm.
8. Detector array dosimetry audit methodology

Figure 8.8 Comparison of the gamma index distribution between the 2D-ARRAY 729 and Gafchromic film in a case where the absolute point dose difference was <1% and a case where the dose difference was 4%. In the former case it can be seen there is good agreement between the 2D-ARRAY 729 and film. In the latter case, the difference between the absolute and relative gamma index is evident.

8.9.3 Timing study

The minimum time taken to perform an audit ranged from a minimum of 6.5 hours to a maximum of 10.9 hours, with a median time of 8.0 hours. Figure 8.9 shows a typical timescale for an audit visit assuming a 9:00 am start time; the median time for each event is plotted. The time taken to perform individual 2D-ARRAY 729 plan measurements had a median of 13 minutes (range 7 – 30 minutes), whereas semiflex ionization chamber and film minutes took a median of 10 minutes (5 – 25 minutes). The additional median time for the 2D-ARRAY 729 are attributed to time taken to perform an online gamma index analysis. The alanine plan measurements accounted for the most amount of time per event, with a median time of 45 minutes (35 – 90 minutes). Other events such as the initial setup, output measurements, 2D-ARRAY 729 cross-calibrations, and setups between measurements had a combined median time of 2.3 hours (1.3 – 4.3 hours).

Using this data it is possible to estimate the likely time taken if only the 2D-ARRAY 729 was used for plan measurements as intended. It is estimated that the median time will be 3.5 hours per audit visit.
As such it is anticipated that most audit visits within the national audit would be possible to be completed in half a working day.

Figure 8.9 Median timescale for audit visits, assuming a 9:00 am start time.

8.10 DISCUSSION

This study demonstrates that, in a direct comparison, the 2D-Array agrees and correlates well with both ion chambers and alanine for dose point measurements. This is shown by the strong concordance correlation coefficients being >0.998 in all cases and this strong trend can be seen in the scatter plots in Figure 8.6. However, the corresponding Bland-Altman plots in Figure 8.6 indicate that there was a systematic tendency for the 2D-Array to slightly under-respond relative to ion chambers and alanine, and that this was independent of the dose level; i.e. the difference was consistent for low to high dose points. However, the average negative difference between the 2D-Array and semiflex ionization chambers and alanine, of approximately -1%, can be mainly attributed to influence from the inhomogeneities caused by the air-filled ion chambers within the array and also to the requirement to use the different OCTAVIUS phantom bases, since the semi-cricoid correction air gap, designed to compensate for the inherent under-response of the 2D-Array posteriorly does not provide a perfect or uniform compensation. This difference is low and within measurement uncertainty [62]. The direct comparison between ion chamber and alanine dose point measurements showed much smaller differences. This agreement is due to the use of the same solid
OCTAVIUS base as described previously and because the calibration of both systems is traceable to the same primary standard.

Of interest is that when the measured dose points are compared with the predicted dose from the TPS it can be seen in the histogram (Figure 8.7a) that the trend is similar for each of the 2D-Array, semiflex ionization chamber and alanine data. When looking at the mean difference between the measured dose and the predicted dose, it was found that the results from the 2D-Array analysis were within 0.5% of those from the semiflex and alanine analysis for measurements in the high dose low gradient region (PTV2 in the 3DTPS test). This shows that the difference between the 2D-Array measurement and the predicted dose calculated in the OCTAVIUS phantom, with the 2D-Array scanned in situ, is similar to the difference found between the semiflex ionization chamber/alanine measurement and the predicted dose calculated in the homogeneous OCTAVIUS phantom scan. This suggests that it is valid to use the 2D-Array data to report individual dose point results in regions of interest to the audited centres, in addition to planar comparisons using the gamma index.

For the gamma index analysis comparison between the 2D-Array and Gafchromic film it was found that the relative comparison of the film in OmniPro masked some of the absolute dose differences due to the normalization procedure described in the methods section; as shown in Figure 8.8. For the 2D-Array measurements it was possible to perform an absolute gamma index analysis against the TPS predicted dose in the Verisoft software. It can be seen in the histograms in Figure 8.7b that more 2D-Array planes analysed using absolute γ would fail compared to film if a passing criteria of ≥95% detector or pixels should pass with a γ<1 for the commonly used passing criteria of 3%/3mm. In general, where there were 2D-Array failures for passing criteria of 3%/3mm, it was found that the absolute dose difference in the point chosen for the gamma index normalisation was close to, or outside, ±3%; however there wasn’t sufficient data to test statistical significant. Practically, the ability of the 2D-Array to give an immediate absolute result was beneficial and allowed for a direct investigation of any unexpected results as they arose during the visits.

Various studies have compared and evaluated detector arrays for routine IMRT and VMAT verification [62,63,66,70,74,75,92,110,113,148,149]. However, at the time of writing, use of these detectors for radiotherapy dosimetry audits has not been reported in the literature and the only report available for a dosimetry audit for rotational radiotherapy was in a Swiss IMRT audit [42] which included RapidArc in the comparison. Dosimetry audits can be effectively performed by using only the 2D-Array only for plan measurements. During each visit, a simple timing study was
performed where the time to perform specific events such as plan measurements was recorded. It was estimated that the time taken if only the 2D-Array was used for a rotational radiotherapy dosimetry audit, the total time per visit would be a median of 3.5 hours (2.2 – 6.7 hours). This streamlined time would allow these audits to be practically performed in an afternoon or evening which will be more favourable for busy radiotherapy centres. Palmer et al. [37] recommend that in order to ensure good compliance rates, audit time should be kept to within four hours of machine time.

Using the methodology developed in this study, a national dosimetry audit of rotational radiotherapy using the 2D-Array and OCTAVIUS phantom only for plan measurements was performed in the United Kingdom [150]. There was a requirement that the audited centre must have started, or be ready to start, clinical treatments with rotational radiotherapy. The audit used the 3DTPS as a baseline test and furthermore offered clinical trial credentialing measurements. As part of the audit, the other systems of dosimetry were taken as back up. The semiflex ionization chambers and alanine continued to be used for standard output measurements.
9

A METHODOLOGY FOR DERIVING OPTIMAL GAMMA INDEX ACCEPTANCE CRITERIA FOR DETECTOR ARRAY MEASUREMENTS

9.1 INTRODUCTION

As discussed previously in this thesis, the gamma index (γ) method is one of the most common techniques for comparing measured and predicted dose distributions [45]. The number of points passing with γ<1 for criteria of 3%/3mm, is the most frequently reported parameter in the literature, with a typical tolerance for the passing rate being 95%. However, the use of the γ passing rate has been shown to have a weak correlation against clinically relevant metrics and the result has also been shown to vary depending on the QA system and software used.

Furthermore the passing rate may not give enough meaningful information, particularly when performing dose delivery benchmarking studies or trend analysis, where subtle differences are of interest. As an example of the passing rate limitation, it is possible to have a systematic dose difference of 2% but the passing rate would still be 100% if 3% acceptance criteria were used for the dose difference. Other metrics that could be extracted from the gamma index map are possible but have not been extensively evaluated, and there are no significant reports in the literature about typical acceptance values.

The purpose of this study was to investigate whether the use of the mean, median, maximum, or near-maximum gamma index (i.e. the maximum gamma index in 1% of points; γ_{1%}) could be suitable alternatives to the passing rate and what acceptance criteria could be used.
9. Methodology for deriving acceptance criteria for detector arrays

9.2 Methodology

9.2.1 Deliberate plan delivery deviation simulations

To evaluate and benchmark different metrics for the gamma index, the 3DTPS test was used with deliberate changes introduced. The expected \( \gamma \) calculations were simulated, in a similar methodology as described in Chapter 4, by exporting the normal plan and perturbed plan predicted dose distributions in DICOM format. The predicted \( \gamma \) was calculated, using in house Matlab-based independent gamma index calculation software, for a 2D coronal & 2D sagittal plane through the geometric centre of the virtual phantom; see Figure 8.3. The 2D perturbed plan planes were compared against a 3D volume for the gamma index calculation. Additionally, a full 3D volume gamma index calculation was performed by comparing the 3D deliberately changes plan DICOM cube against the 3D normal plan DICOM cube. The \( \gamma \) index matrices were interrogated to give passing rates as well as mean, median, max and near-max \( \gamma \) (\( \gamma_{1\%} \)) metrics. A global gamma index passing criterion of 3%/2mm with 20% threshold was used for these investigations as a starting point.

9.2.1.1 Dose deviations

Systematic dose deviations will be mainly attributed to inaccurate TPS modelling parameters related to the modelling of the dynamic MLC motion. To simulate this, purposeful dose modifications were introduced ranging from -5% to +5% in 1% increments. These were performed by offsetting the normalization value of the plans within the Eclipse TPS. This subsequently resulted in a MU offset by the modified amount. It should be noted that the increase in MU also affects the entire MLC leaf bank motion.

9.2.1.2 Spatial deviations

Spatial deviations were simulated with deliberate changes of 1, 2, 3, 4, 5mm throughout all control points for one 5mm MLC. Alignment errors were not simulated. The justification for excluding these is illustrated in Figure 9.1, where it can be seen that these types of errors cause significant failures and in a real-scenario the measurement setup would be checked and a diagnosis of incorrect setup would be rectified by repeating the measurement. As such, the final result that would be recorded would be one after a repeat measurement was taken. For the above dose and spatial deviations, a similar scenario would occur whereby the setup would be checked, calibrations would be repeated,
and measurements repeated as necessary, but there is a higher likelihood that they would be found to be consistent and therefore be recorded in the final results.

9.2.1.3 Combined dose and spatial deviations

The dose deviations and the spatial deviations described above were combined to represent a more realistic scenario. For each MLC deliberate change plan, a dosimetric deviation of +3% and -3% dose was introduced.

![Figure 9.1 Example gamma index distribution for a 3mm Left-right alignment deviation. Calculated using global gamma 3%/2mm.](image)

9.2.2 Retrospective analysis of measured data from the UK national audit of VMAT and Tomotherapy

Following on from the simulation study, the data from the completed UK national audit of VMAT and Tomotherapy [150] was used to investigate whether similar trends could be observed. The audit was based on the methodology described in Chapter 7. The Matlab-based software was used to retrospectively re-analyse 80 dose planes measured for the 3DTPS virtual phantom test in the coronal 1 and sagittal plane. The software was set up to automatically calculate global $\gamma$ calculations for 3%/2mm for a 20% threshold and in line with the simulation study, the $\gamma$ passing rate, $\gamma_{\text{mean}}$, $\gamma_{\text{median}}$, $\gamma_{\text{max}}$, and $\gamma_{1\%}$ were calculated.
9. Methodology for deriving acceptance criteria for detector arrays

9.2.3 Correlation of gamma index metrics against treatment plan DVH metrics

To assess the suitability of the different gamma index metrics; the correlation against treatment plan DVH metrics was assessed. Previous studies have attempted to correlate the $\gamma$ passing rates against DVH metrics and have reported weak correlations \[86,151,152\]. At the time of writing, no reports in the literature could be found which have investigated whether it is possible to correlate other metrics such as the mean or median calculated $\gamma$ against DVH metrics.

9.2.3.1 Correlation based on simulation data

As a feasibility study these values were calculated for the 3DTPS test for a 2D coronal & sagittal plane similar to Figure 8.3 and for a 3D volume gamma calculation. The primary PTV (PTV2) mean dose and the OAR max dose for the 3DTPS test were calculated for the normal plan as well as the perturbed plans. The percentage difference between the perturbed plan DVH metrics and the normal plan were calculated to use in the correlation analysis; which made use of the Pearson correlation coefficient. Due to the absolute nature of the gamma index calculation it would not be possible to correlate against DVH metrics that had a range of positive and negative values. Therefore the DVH metric percentage differences were calculated as absolute % dose difference:

$$\Delta D\% = \left| \left( \frac{Dose_{error} - Dose_{normal}}{Dose_{normal}} \right) \times 100\% \right|$$ \[9.1\]

9.2.4 Correlation based on measured data

Since the 3DTPS plan measurements at all centres taking part in the audit followed the same procedure, it was possible to define regions of interest within the 2D measured planes that corresponded to the different structures of the 3DTPS test. In the coronal plane and the sagittal plane it was possible to define a 2D ROI for PTV2, and in the sagittal plane it was possible to define a 2D ROI for the OAR. As such this made it possible to calculate the mean absolute percentage dose difference within the PTV2 ROI and the maximum absolute percentage dose difference within the OAR ROI as defined using equations 9.2 and 9.3 respectively:

$$\Delta D\%_{PTV2\text{mean}} = \left| \left( \frac{ROI_{PTV2\text{mean,measured}} - ROI_{PTV2\text{mean,TPS}}}{ROI_{PTV2\text{mean,TPS}}} \right) \times 100\% \right|$$ \[9.2\]

$$\Delta D\%_{OAR_{\text{max}}} = \left| \left( \frac{ROI_{OAR_{\text{max,measured}} - ROI_{OAR_{\text{max,TPS}}}}}{ROI_{OAR_{\text{max,TPS}}}} \right) \times 100\% \right|$$ \[9.3\]
9. Methodology for deriving acceptance criteria for detector arrays

9.2.5 Methods for Estimating Gamma Index Acceptance Criteria

After evaluating whether the gamma index metrics could be correlated against DVH metrics, the next step was to assess whether it is possible to estimate acceptance criteria for the different metrics. This was performed using the two methods as discussed below in sections 9.2.5.1 and 9.2.5.2.

9.2.5.1 Receiver Operator Characteristic (ROC) Analysis

One method that has been used to estimate optimal QA acceptance criteria is the Receiver Operator Characteristic (ROC) [153–155]. Using the ROC method, one can evaluate the sensitivity and specificity of a range of acceptance criteria to estimate the most appropriate criterion. This analysis was performed on the simulation study data initially. To be able to perform a ROC analysis, a cut-off value for a true positive should be defined. In this particular case the test is whether the gamma index analysis is able to detect an out-of-tolerance $\Delta D\%$. Therefore the analysis has been performed comparing the $\gamma$ index metrics against the PTV2 mean $\Delta D\%$ and OAR maximum $\Delta D\%$. The threshold value for an out-of-tolerance $\Delta D\%$ has been set at $\pm 3\%$ [28]. This analysis was performed using the simulation study and repeated using the measured data. In order to calculate the ROC curves for each $\gamma$ metric, the sensitivity and specificity were calculated using the following equations:

\[
\text{Sensitivity} = \frac{TP}{TP + FN} \quad (9.4)
\]
\[
\text{Specificity} = \frac{TN}{FP + TN} \quad (9.55)
\]

Where TP, FN, FP and TN are defined as follows:

- **True-positive (TP):** instance where the $\gamma$ metric is out of tolerance and where $\Delta D\%$ is also out of tolerance.
- **False-positive (FP):** instance where the $\gamma$ metric is in tolerance, whereas $\Delta D\%$ is out of tolerance.
- **True-positive (TP):** instance where the $\gamma$ metric is in tolerance and where $\Delta D\%$ is also in tolerance.
- **False-negative (FN):** instance where the $\gamma$ metric is out of tolerance, whereas $\Delta D\%$ is in tolerance.

An ideal metric would have 100% sensitivity (i.e. no false positive results) and 100% specificity (i.e. no false negative results). In practice, however, the optimal acceptance criterion is the one with the best trade-off between sensitivity and specificity. For example, consider two possible criteria on a
9. Methodology for deriving acceptance criteria for detector arrays

ROC curve. The first criterion is calculated with 75% sensitivity and 100% specificity, whereas the second criterion has an associated sensitivity of 100% and specificity of 70%. In the first example, the criterion would mean that potentially 25% of cases that should have failed may be falsely passed. The only advantage is that there is minimum risk that plans that should pass would be recorded as failed by the criterion; meaning fewer resources would be used investigating failed results. This criterion would only be valid if one knows the potential clinical impact of a result that should have failed, but passed based on the criterion. The second criterion would reduce the risk of false positives; however it would consequently mean that a higher proportion of cases may be failed when they should have passed. In a benchmarking setting, subtle differences are critical and therefore an acceptance criterion with 100% sensitivity will be optimal. Therefore acceptance criteria in this study were estimated assuming 100% sensitivity required.

To avoid recommending a gamma index metric that is not statistically reliable, the area under the ROC curve (AUC) was calculated for each $\gamma$ metric. The number varies from 0.5 – 1. For all of the ROC calculations above, the AUC was also calculated. The following guide has been used to interpret the results [156]:

- 1 = Perfect test
- $0.9 < \text{AUC} < 1$ = highly accurate
- $0.7 < \text{AUC} \leq 0.9$ = moderately accurate
- $0.5 < \text{AUC} \leq 0.7$ = less accurate
- $0.5$ = non-informative

The acceptance criterion for each $\gamma$ metric was estimated assuming 100% sensitivity required, with the associated specificity recorded. The AUC was used to estimate whether that metric is a reliable test.

9.2.5.2 Estimating Acceptance Criteria Based on AAPM TG119 Recommendations

Another method is to use measured data to retrospectively estimate acceptance criteria based on statistical confidence limit (CL) calculations. This was originally proposed by Venselaar et al who use the following equation to calculate acceptance criteria for dose difference, $\delta_{\Delta D}$ [157]:

$$\delta_{\Delta D} = \sqrt{\langle \Delta D \rangle} + 1.5\sigma$$  \hfill (9.7)
where \( \sigma \) is standard deviation.

The AAPM TG119 report made use of confidence interval calculations to specify acceptance criteria for the percentage dose difference between calculated and planned doses and for the gamma index passing rate for 3%/3mm [139]. The report calculated the 95% confidence interval by substituting the 1.5 multiplier in equation (9.7) by 1.96. Additionally, the following equation was proposed to calculate acceptance criteria for gamma index passing rate, \( \delta_\Gamma \) [139]:

\[
\delta_\Gamma = 100 - \left( \Gamma_{\text{mean}} + 1.96\sigma \right)
\]  

(9.8)

where \( \Gamma \) is the gamma index passing rate.

The choice of multiplication factor will determine how strict the calculated acceptance criteria. In a similar analogy to the choice of sensitivity and specificity target values when performing ROC analysis, as described in section 9.2.5.1., the choice of formalism for calculating the CI is open to interpretation depending on requirements. Equation (9.7) was used to calculate acceptance criteria for the \( \gamma_{\text{mean}} \), \( \gamma_{\text{median}} \), \( \gamma_{\text{max}} \), and \( \gamma_{1\%} \) and equation (9.8) was used to calculate acceptance criterion for passing rate to compare against those calculated by ROC analysis. In this study, calculations were performed with multiplication factors of 1.5 and 1.96 for comparison.

This method for estimating acceptance criteria is reliant on the data available having a Gaussian distribution and on the data having no significant outliers which would skew the CI calculation and thus result in more lenient acceptance criteria. For these reasons it was not possible to use this calculation on the simulated deliberate modification plans.

9.2.6 Comparison of different passing criteria

All previous evaluations above have been made for passing criteria of 3%/2mm. It is important to understand whether this methodology would be transferrable to other systems. As has been shown in Chapter 6, different systems could report different results for the same passing criteria. However, it was also shown that passing criteria from one system could be modified to give agreeable results with the results from another system using different passing criteria. For example 4%/3mm on the Delta4 phantom gave results which agreed with 3%/3mm on the PTW 2D-ARRAY 729. Varying the passing criteria in this study and correlating against 3%/2mm could be used as a surrogate to test the robustness of this methodology. For the simulation study and the retrospective analysis, the gamma index was re-calculated for passing criteria of 3%/3mm and 2%/2mm for a threshold of 20%. These
were performed to check the consistency of the trend of the gamma index results against 3%/2mm and therefore to understand whether the passing thresholds could be scaled for different passing criteria. Comparisons were carried out for the $\gamma$ passing rate and $\gamma_{\text{mean}}$.

9.2.7 Evaluation of different lower thresholds for the gamma index

It is common to limit the gamma index calculation to all points that are $\geq$10-20% of the maximum dose value within the dose distribution [49]. This is to eliminate dose in the out-of-field region where a large relative dose difference can be calculated and skew the gamma index result (whereas the absolute dose difference is small relative to the prescription point). However, it was interesting to evaluate different threshold values other than 20% to check whether this value is appropriate, or whether a lower/higher value should be used. Therefore the analysis of the NRRA data was repeated by varying the lower threshold from 0% up to 90%. Calculations were made for global and local gamma. For each threshold the mean gamma index and passing rate was calculated.

9.3 RESULTS

9.3.1 Correlation of gamma index metrics against DVH metrics

9.3.1.1 Simulation Study

The $\gamma_{\text{mean}}$, $\gamma_{\text{median}}$ and $\gamma_{1\%}$ metrics had very strong statistically significant correlations against the PTV2 mean DVH metric for 3D volumetric $\gamma$ calculation as well as 2D planar $\gamma$ ($p>0.95$, $p<0.01$); this can be seen in the data in Table 9.1. Figure 9.2 shows the trend of $\gamma$ metrics against the absolute PTV2 mean dose deviation, and Figure 9.3 shows the correlation against the absolute OAR maximum dose deviation. The $\gamma$ passing rate had a comparatively weaker linear correlation against the PTV2 DVH metrics; however a non-linear trend was apparent. The maximum $\gamma$ had poor correlation against the PTV2 metrics. The opposite trend was found with the correlation of $\gamma$ metrics against the OAR DVH metric ($p>0.95$, $p<0.01$). The maximum $\gamma$ had a strong correlation, whereas the remaining metrics suffered from poor correlation.
9. Methodology for deriving acceptance criteria for detector arrays

Figure 9.2 3%/2mm global 3D volume (a) gamma index pass-rate, (b) mean gamma, (c) median gamma, (d) max gamma, (e) near-max gamma against the PTV2 mean absolute dose deviation.
9. Methodology for deriving acceptance criteria for detector arrays

Figure 9.3 3%/2mm global 3D volume (a) pass-rate, (b) mean gamma, (c) median gamma, (d) max gamma, (e) near-max gamma against OAR maximum dose deviation.
Table 9.1 Summary of mean, coefficient of variance (c_v) and Pearson correlation coefficient (ρ) for each γ parameter for 3%/3mm, 3%/2mm, 2%/2mm using a 20% threshold. All ρ have p<0.01.

<table>
<thead>
<tr>
<th>γ metric</th>
<th>DVH metric</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>PTV2 mean</td>
<td>OAR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>p</td>
<td>p-value</td>
<td>p</td>
<td>p-value</td>
</tr>
<tr>
<td>3D volume gamma</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passing rate</td>
<td>-0.76</td>
<td>&lt;0.01</td>
<td>-0.07</td>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td>γ mean</td>
<td>0.97</td>
<td>&lt;0.01</td>
<td>0.02</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>γ median</td>
<td>0.95</td>
<td>&lt;0.01</td>
<td>-0.03</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>γ max</td>
<td>0.19</td>
<td>0.36</td>
<td>0.93</td>
<td>&lt;0.01</td>
<td></td>
</tr>
<tr>
<td>γ 1%</td>
<td>0.96</td>
<td>&lt;0.01</td>
<td>0.22</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>2D Coronal Plane</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passing rate</td>
<td>-0.81</td>
<td>&lt;0.01</td>
<td>-0.15</td>
<td></td>
<td>0.47</td>
</tr>
<tr>
<td>γ mean</td>
<td>0.97</td>
<td>&lt;0.01</td>
<td>0.04</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td>γ median</td>
<td>0.95</td>
<td>&lt;0.01</td>
<td>-0.03</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>γ max</td>
<td>0.49</td>
<td>0.01</td>
<td>0.82</td>
<td>&lt;0.01</td>
<td></td>
</tr>
<tr>
<td>γ 1%</td>
<td>0.88</td>
<td>&lt;0.01</td>
<td>0.44</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>2D Sagittal Plane</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passing rate</td>
<td>-0.77</td>
<td>&lt;0.01</td>
<td>-0.05</td>
<td></td>
<td>0.81</td>
</tr>
<tr>
<td>γ mean</td>
<td>0.96</td>
<td>&lt;0.01</td>
<td>0.01</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>γ median</td>
<td>0.94</td>
<td>&lt;0.01</td>
<td>-0.03</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>γ max</td>
<td>0.17</td>
<td>0.41</td>
<td>0.92</td>
<td>&lt;0.01</td>
<td></td>
</tr>
<tr>
<td>γ 1%</td>
<td>0.90</td>
<td>&lt;0.01</td>
<td>0.44</td>
<td>0.03</td>
<td></td>
</tr>
</tbody>
</table>
9. Methodology for deriving acceptance criteria for detector arrays

Figure 9.4 shows the correlation between a 3D volumetric gamma index analysis vs. analysis made in the 2D coronal and sagittal planes for 3%/2mm and 20% lower dose threshold. It was found that the 2D analysis had statistically strong Pearson correlations of $>0.99$ ($p<0.01$) against 3D for the mean gamma index metric. The passing rate had correlation $>0.95$ ($p<0.01$). The coronal plane had closer agreement with the 3D gamma index.

![Simulated 3D volumetric gamma index vs 2D coronal and sagittal planes for mean gamma index and passing rate for 3%/2mm, 20% threshold.](image)

9.3.1.2 Measured data

The correlation between the $\gamma$ metrics and estimated PTV2 absolute mean dose deviation and OAR absolute maximum dose deviation are shown in Table 9.2 and in Figure 9.5. Lower correlations were found compared to the simulation study but were statistically strong at $\rho > 0.7$ ($p$-value $< 0.05$ in all cases).

<table>
<thead>
<tr>
<th>$\gamma$ metric</th>
<th>DVH metric</th>
<th>PTV2 mean</th>
<th>OAR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho$</td>
<td>$p$-value</td>
<td>$\rho$</td>
</tr>
<tr>
<td>Passing rate</td>
<td>-0.78</td>
<td>$&lt;0.01$</td>
<td>-0.64</td>
</tr>
<tr>
<td>$\gamma_{\text{mean}}$</td>
<td>0.74</td>
<td>$&lt;0.01$</td>
<td>0.58</td>
</tr>
<tr>
<td>$\gamma_{\text{median}}$</td>
<td>0.69</td>
<td>$&lt;0.01$</td>
<td>0.51</td>
</tr>
<tr>
<td>$\gamma_{\text{max}}$</td>
<td>0.75</td>
<td>$&lt;0.01$</td>
<td>0.65</td>
</tr>
<tr>
<td>$\gamma_{1%}$</td>
<td>0.78</td>
<td>$&lt;0.01$</td>
<td>0.69</td>
</tr>
</tbody>
</table>
Figure 9.5  3%/2mm, 20% threshold global gamma index passing rate (points passing with $\gamma<1$) given as a percentage (plotted on the left $y$-axis), median, mean and near-max gamma index (plotted on the right $y$-axis) plotted against (a) the mean absolute dose difference in the 2D ROI corresponding to PTV2 mean and (b) the absolute maximum dose difference in the 2D ROI corresponding to OAR. The trend line for the $\gamma_{\text{max}}$ is not displayed.
9. Methodology for deriving acceptance criteria for detector arrays

9.3.2 Acceptance criteria based on ROC analysis

9.3.2.1 Derivation from simulation study

The ROC curves for the gamma index metrics accuracy in detecting outside of ±3% dose deviation in the PTV2 mean and OAR maximum dose DVH metrics are shown in Figure 9.6a and Figure 9.6b, and for the measured data shown in Figure 9.6c and Figure 9.6d. The diagonal line indicates a random guess and also equates to AUC=0.5. The estimated acceptance criteria, along with the specificity and AUC are given in Table 9.3. The AUC values are given with their corresponding p-value where statistical significance is taken as p<0.05.

In the simulation study, the gamma index metrics, except for the γ_{max} had an AUC indicating high accuracy in detecting dose deviations in PTV2. The γ_{max} metric was close to the diagonal line on the ROC curve and therefore is statistically not reliable. A similar trend was seen in the measured data, except that the γ_{max} had a better AUC.

The γ metrics were moderately accurate against the OAR max DVH metric according to the AUC values; however, in order to achieve 100% for these metrics, the acceptance criteria for the passing rate would be required to be 100% and to reduce significantly to <0.2 for the other metrics; with a corresponding low specificity. The γ_{max} metric had the best performance with a statistically highly accurate result of 0.94 (p<0.01) for the AUC.
9. Methodology for deriving acceptance criteria for detector arrays

Figure 9.6  Simulation study ROC analysis of the sensitivity and specificity of gamma metrics to predict a 3% deviation in (top) PTV2 mean dose deviation and (bottom) OAR maximum dose deviation. Gamma analysis performed using 3%/2mm with 20% threshold.
9. Methodology for deriving acceptance criteria for detector arrays

Table 9.3  Acceptance criteria for the gamma index metrics to predict a ±3% dose deviation in the PTV2 mean dose and OAR maximum dose based on the simulation and measured data. In all cases, sensitivity was 100%.

<table>
<thead>
<tr>
<th>y metric</th>
<th>PTV2 mean</th>
<th>DVH metric</th>
<th>OAR max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Acceptance criterion</td>
<td>AUC</td>
<td>Specificity</td>
</tr>
<tr>
<td><strong>Passing rate</strong></td>
<td>≥99.9%</td>
<td>0.95 (p&lt;0.01)</td>
<td>73.3%</td>
</tr>
<tr>
<td>$\gamma_{\text{mean}}$</td>
<td>&lt;0.38</td>
<td>0.96 (p&lt;0.01)</td>
<td>80%</td>
</tr>
<tr>
<td>$\gamma_{\text{median}}$</td>
<td>&lt;0.35</td>
<td>0.97 (p&lt;0.01)</td>
<td>80%</td>
</tr>
<tr>
<td>$\gamma_{\text{max}}$</td>
<td>&lt;0.90</td>
<td>0.64 (p=0.22)</td>
<td>33.3</td>
</tr>
<tr>
<td>$\gamma_{1%}$</td>
<td>&lt;0.85</td>
<td>0.96 (p&lt;0.01)</td>
<td>80%</td>
</tr>
</tbody>
</table>

Table 9.4  Acceptance criteria based on confidence interval calculations for the measured audit data for Global 3%/2mm, 20% threshold.

<table>
<thead>
<tr>
<th>C.L. Formalism</th>
<th>Metric</th>
<th>Pass-rate</th>
<th>$\gamma_{\text{mean}}$</th>
<th>$\gamma_{\text{median}}$</th>
<th>$\gamma_{\text{max}}$</th>
<th>$\gamma_{1%}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAPM TG119 [139]</td>
<td>≥ 92.3%</td>
<td>&lt;0.55</td>
<td>&lt;0.54</td>
<td>&lt;1.73</td>
<td>&lt;1.28</td>
<td></td>
</tr>
<tr>
<td>Venselaar et al [157]</td>
<td>≥ 93.7%</td>
<td>&lt;0.50</td>
<td>&lt;0.49</td>
<td>&lt;1.58</td>
<td>&lt;1.18</td>
<td></td>
</tr>
</tbody>
</table>

9.3.3  Acceptance Criteria Based on Confidence Limits

The estimated acceptance criteria based on calculating confidence intervals for the different global gamma index metrics are given in Table 9.4 for passing criteria of 3%/2mm and a lower dose threshold of 20%.

Table 9.4  Acceptance criteria based on confidence interval calculations for the measured audit data for Global 3%/2mm, 20% threshold.

<table>
<thead>
<tr>
<th>C.L. Formalism</th>
<th>Metric</th>
<th>Pass-rate</th>
<th>$\gamma_{\text{mean}}$</th>
<th>$\gamma_{\text{median}}$</th>
<th>$\gamma_{\text{max}}$</th>
<th>$\gamma_{1%}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAPM TG119 [139]</td>
<td>≥ 92.3%</td>
<td>&lt;0.55</td>
<td>&lt;0.54</td>
<td>&lt;1.73</td>
<td>&lt;1.28</td>
<td></td>
</tr>
<tr>
<td>Venselaar et al [157]</td>
<td>≥ 93.7%</td>
<td>&lt;0.50</td>
<td>&lt;0.49</td>
<td>&lt;1.58</td>
<td>&lt;1.18</td>
<td></td>
</tr>
</tbody>
</table>

9.3.4  Correlation of different gamma index passing criteria

The correlation of passing criteria 3%/3mm, 2%/2mm, and 2%/3mm with 3%/2mm for the mean gamma index metric and passing rate is given in Figure 9.7a and b for the simulation study and in
Figure 9.7c and d for the measured audit data. The mean gamma correlations were all >0.99 (p<0.01) in both cases. The passing rate had a lower correlation but statistically strong at >0.95 (p<0.01).

Figure 9.7 Simulated (a) and measured (b) mean gamma index values for passing criteria 3%/3mm and 2%/2mm against 3%/2mm. The dashed line indicates a 1:1 agreement.

9.3.5 The effect of varying the lower dose threshold for γ analysis

The impact of varying the lower dose threshold on the gamma index analysis is shown in Figure 9.8 for the global mean gamma index metric and passing rate. The results show a non-linear trend which is similar between local and global gamma index. It is seen that there is a steep variation in the
9. Methodology for deriving acceptance criteria for detector arrays

gamma index result from a lower dose threshold of 0% up 10% for global gamma and up to 5% for local gamma. After this point the trend becomes more stable.

![Figure 9.8](image)

Figure 9.8 Average (a) $\gamma_{\text{mean}}$ (b) gamma index passing rate for the measured data as a function of lower dose threshold. The error bars represent standard error of the mean. The line linking the individual points is for visualization only.

9.4 DISCUSSION

It has been shown that useful information may be extracted from gamma index calculations that have not been extensively studied in the literature. Figure 9.2 and Figure 9.3 demonstrate the limitations of using the passing rate in correlations against DVH metrics. As the passing rate only evaluates whether points are above or below $\gamma=1$, there becomes a threshold below which points will always be $\gamma<1$. For example it is possible when using 3%/2mm passing criteria, that a systematic dose difference of $\pm2\%$ the measurement would still pass as the pass-rate would be 100%. As shown in Figure 9.2 and Figure 9.3, the trend line represents the best fit using the least squares method, however it would be possible, depending on the algorithm used, to fit a range of different linear trends, and therefore this should only be taken as a guide. Hence the passing rate has weak statistical robustness in predicting the impact on DVH metrics. The other gamma index metrics could be more robustly correlated as they have no upper limit on their value. In the simulated and the measured data, the $\gamma_{\text{mean}}, \gamma_{\text{median}}$ and $\gamma_{1\%}$ demonstrated statistically strong correlations with the PTV2 mean dose DVH metric. However, they suffered from poor correlation against the OAR DVH metric. Interestingly, this weak correlation was not as apparent in the measurement data as the simulated data, possibly due to the resolution of the detector array smoothing out the maximum point difference. The $\gamma_{\text{max}}$ correlated well with the OAR metric but had a poor correlation against PTV2 in
9. Methodology for deriving acceptance criteria for detector arrays

the simulation study. Hence is important to note that no single γ metric will give all the required information for both PTVs and OARs.

The acceptance criteria for the different γ metrics based on the ROC analysis varied between the simulation study and the measured data and also between the prediction of an outside ±3% deviation in the PTV2 mean dose DVH metric and OAR maximum dose DVH metric. The \( \gamma_{\text{mean}} \) and \( \gamma_{\text{median}} \) metrics showed the best consistency for predicting a PTV2 mean dose deviation for 3%/2mm and a lower dose threshold of 20% across the simulated and measured data. Using the methods for calculating acceptance criteria based on the confidence limits gave larger numbers than those derived based on ROC analysis but there was overall closer agreement between the values for \( \gamma_{\text{mean}} \). The main disadvantage of the confidence interval approach is that it is not directly related to a given dose deviation and also relies on having good quality measured data to avoid skewing the criteria estimate. For this reason, using ROC analysis is statistically more robust for deriving acceptance criteria. Using the calculations from the different methods, it appears that an acceptance value of 0.45 for the \( \gamma_{\text{mean}} \) would be appropriate for 3%/2mm.

By varying the dose deviation variable in the ROC analysis between 1% – 5% it was possible to estimate guideline criterion for different dose deviations. For each dose deviation variable, the optimal acceptance criterion was estimated assuming 100% sensitivity. Table 9.5 shows guideline criterion for \( \gamma_{\text{mean}} \) for this study. This can be supplemented by also evaluating the maximum gamma index value with particularly the simulation study analysis indicating that points with \( \gamma > 1.1 \) may be associated with a >3% increase in the OAR maximum dose.

<table>
<thead>
<tr>
<th>Mean γ criterion</th>
<th>Indicative dose deviation</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.15</td>
<td>±1%</td>
<td>0.75</td>
</tr>
<tr>
<td>0.30</td>
<td>±2%</td>
<td>0.83</td>
</tr>
<tr>
<td>0.45</td>
<td>±3%</td>
<td>0.93</td>
</tr>
<tr>
<td>0.60</td>
<td>±4%</td>
<td>0.97</td>
</tr>
<tr>
<td>0.90</td>
<td>±5%</td>
<td>1.00</td>
</tr>
</tbody>
</table>

As a proof-of-concept, 30 Head & Neck cases that were also measured during the national audit were analysed using ROC. The endpoint of ±3% primary PTV dose deviation was tested for the global
mean gamma criterion (for 3%/2mm, 20% threshold). The optimal acceptance criterion was found to be 0.43 (sensitivity 100%, specificity 62%, AUC=0.750, \( p=0.04 \)) in good agreement with the value derived from the 3DTPS test for predicting a primary PTV dose deviation. This indicates that the values derived in this study are valid for the PTW 2D-ARRAY 729 and OCTAVIUS combination rather than being plan-specific.

The simulation acceptance criteria were derived based on a 3D volumetric gamma index calculation. There was a strong correlation between the 3D \( \gamma \) volume metrics and the 2D coronal plane and sagittal plane \( \gamma \) calculations. The coronal plane \( \gamma \) had a closer agreement to the 3D volume \( \gamma \), with a bigger difference seen in the sagittal plane. This highlights that the configuration of the measurement plane should be characterized to check whether different criterion would be needed. To date, this has also only been tested on PTW planar 2D-Array measured data. As in Chapter 6, different software and hardware combinations could give different gamma index results, and it would therefore be necessary to perform this evaluation on the other commercial systems to be able to propose universal (or manufacturer specific) tolerances. However, the aim of this study was to develop a framework for investigating different metrics and outlining a methodology for arriving at optimal acceptance values. This can be supported by Figure 9.7a and Figure 9.7b, where there is evidence that the \( \gamma_{\text{mean}} \) varies linearly with changing passing criteria and so, theoretically, a similar linear correlation would be expected if the data were measured using a different system and compared with results in this study. There is a strong correlation (\( \rho>0.99, p<0.01 \)) between mean gamma calculated for 3%/2mm against other passing criteria, which also shows statistical robustness of the \( \gamma_{\text{mean}} \). As shown in Figure 9.7c and Figure 9.7d, the passing criteria shows more dispersion of the data although the correlation is still strong at \( \rho>0.95 \).

The impact of varying the lower dose threshold for the gamma index calculation has been investigated. It can be seen in Figure 9.8a and Figure 9.8b respectively, that \( \gamma_{\text{mean}} \) and passing-rate vary non-linearly with changing the lower dose threshold. For global \( \gamma_{\text{mean}} \) calculations, there was trend towards a very steep increase from 0 – 10% threshold, which changes to a lower gradient linear increase between 10 – 40% threshold, and then gradually reducing in gradient >40% until plateauing above 80%. A similar inverse trend is seen in the passing rate, except that the non-linearity became more pronounced at >20% rather than 40%. This indicates that the choice of an acceptance tolerance value must always be specified with a fixed lower threshold value used. Additionally the choice of the lower threshold should be chosen in a region of the trend where there is low gradient to ensure the tolerance value is robust to slight variations in measurement.
uncertainty. In particular the acceptance criteria should not be specified for a threshold <10%. A value of 20% is therefore considered suitably robust.

This methodology can be further extended by measuring the deliberately changed plans using 2D and 3D detector array/phantom measurements to test the criteria estimated by the simulation study. In a benchmark setting, this methodology could be used prospectively by introducing the same types of deliberate changes into the test plan(s) to be used and measuring those plans using the equipment that will be used to determine suitable acceptance criteria.

In conclusion, the $\gamma$ passing rate may lack statistical robustness to be suitable for benchmarking of dose delivery. It is more suited to routine clinical measurements, once there is confidence in the accuracy of the TPS and linac hardware commissioning. Van Esch et al [158] give a pragmatic approach for sensible usage of the passing rate in this scenario. The $\gamma_{\text{mean}}$ metric has better potential to be used as a parameter to predict PTV dose deviation as the value is not bound by an upper limit and can give a more useful guide regarding potential deviations in dose difference. However this needs to be complemented by the use of the $\gamma_{\text{max}}$ for OAR dose deviations. Future work should be to assess whether the $\gamma_{\text{mean}}$ or $\gamma_{\text{median}}$ metrics could be used to correlate against machine-specific parameters such as the MU or against metrics that score the complexity of a particular plan [159,160].
10

DISCUSSION AND CONCLUSIONS

10.1 DISCUSSION

This thesis has focussed on evaluating the performance of detector arrays for the verification of advanced IMRT and VMAT treatment plans. Commercial detector arrays with different geometrical configurations to one another in attempting to find an optimal, yet practical, solution to verification of intensity modulated radiotherapy. Through development of characterisation tests, it has been possible to assess the response of various commercial detector array systems and their combined gamma index analysis software. Deliberately inserted changes were designed to test the QA systems, however, the likelihood of this type of deviation occurring in clinical practice should be considered. An MLC motor is affected due to wear-and-tear, leading to a leaf travelling slower than expected and therefore the leaf lags behind the other leaves, in a way that would be similar to the modifications simulated in this thesis. The tolerance on the Linac MLC control software (commonly 2mm on a Varian linac), means that there are generally two possible feedback scenarios: if possible, all the other leaves are slowed down and the dose rate is decreased to compensate for the slower leaf; or an interlock may be activated. Software errors may also lead to a mis-translation of the MLC positions. This study has demonstrated that in the event of subtle plan deviations the detector array systems evaluated are able to detect some of these if suitable gamma index passing criteria, such as 2%/2mm, are used. However, even lower passing criteria may be required for film, EPID, and ArcCHECK. Heilemann et al also concluded that 2%/2mm is necessary to detect positional deviations in MLCs for RapidArc deliveries [76]. This work agrees well with various other studies that have been performed on the impact of errors in different detector array systems [75,85,92,110,113,149].

The common factor in all commercial array systems is the reliance on the gamma index method to provide the quantitative evaluation of the measured dose distribution against the TPS calculated dose distribution. The mathematical definition of the gamma index is straightforward; however it has presented challenges from the computing sense in terms of the speed of the calculation of the metric. The different computational approaches that are possible can produce variability in the calculation of the gamma index between different software. This has been demonstrated in Chapter...
5 where a bespoke Matlab software was designed with the flexibility to vary different parameters. Two of these parameters were (a) whether or not to interpolate the evaluated dose distribution such that the pixel spacing is sufficiently less than the distance criterion to avoid uncertainties in the gamma calculation, and (b) whether to search the entire evaluated distribution or limit the search. It has been shown that for two commercial software (OmniPro ImRT 7.0 and Verisoft 5.1) that one performed such an interpolation whilst the other did not. The results in this thesis indicate that it is important to understand the response and limitations of the gamma index analysis combined with the hardware/software equipment in use. For the same passing criteria, different devices and software combinations exhibit varying levels of agreement with the Matlab predicted gamma index analysis. As shown in this study, passing criteria of say 3%/3mm may not give the same results for measurements by different QA systems.

It has been shown that the impact of the 2D and 3D gamma index can be significant depending on the spacing of detectors within an array and the way that it is used. In the case of a planar measurement using the sparsely arranged 2D-ARRAY 729 (as originally intended for this piece of equipment) the gamma index passing rate was in reasonable agreement with the predicted 3D volumetric passing rate. However when the 2D-ARRAY 729 was used in the OCTAVIUS 4D phantom to reconstruct a 3D dose distribution, it gave the worst overall agreement at $\rho_c = 0.28$.

The suitability of the gamma index evaluation method in detecting clinically significant deviations has previously been questioned [86] and alternatives have been suggested [86,161–163]. However, the gamma index has been widely accepted and is implemented into most commercial software. The gamma index provides the means for an efficient analysis which is particularly important within a busy clinical environment [164]. It has also been used effectively within dosimetry audits of complex radiotherapy [41,43,106,150]. If one is performing a retrospective analysis of patient-specific QA in order to streamline the process, the gamma index provides a suitable means to explore trends over a period of time [165,166].

This work has found that it is suitable to use a detector array in a dosimetry audit of rotational radiotherapy in place of standard systems of dosimetry such as ion chambers, alanine and film. In a direct comparison, the 2D-Array agrees and correlates well with both ion chambers and alanine for dose point measurements. This was shown by the strong concordance correlation coefficients being >0.998 in all cases. For the gamma index analysis comparison between the 2D-Array and Gafchromic film it was found that the relative comparison of the film in OmniPro masked some of the absolute
dose differences due to the normalization procedure described in the methods section. For the 2D-
Array measurements it was possible to perform an absolute gamma index analysis against the TPS
predicted dose in the Verisoft software. Practically, the ability of the 2D-Array to give an immediate
absolute result was beneficial and allowed for a direct investigation of any unexpected results as
they arose during the visits. The methodology developed in Chapter 8 has been successfully used in
a UK audit of 34 cancer centres with 43 treatment delivery systems [150].

A follow-up analysis of the library of measured data during the audit found that additional metrics
such as the mean gamma index or dose differences over regions of interest can be gleaned from the
measured dose distributions as demonstrated in Chapter 9. A methodology for being able to
prospectively ascertain appropriate gamma index acceptance criteria for the detector array system
in use, via simulation of deliberate plan changes and ROC analysis, has been developed. It has been
shown that setting appropriate tolerances can be achieved and should be performed as the
methodology takes into account the configuration of the commercial system as well as the software
implementation of the gamma index.

10.2 CONCLUSIONS

This thesis has found that there are variable factors affecting the performance of detector arrays in
the accurate verification of IMRT and VMAT treatment plans. These range from the configuration of
the detector array to the software that is used along with its implementation of the gamma index
analysis. In measuring treatment plans with deliberately introduced modifications, it was found that
the different commercial systems (which had different resolution and configuration) were able to
detect clinically relevant deviations as long as the correct gamma index passing criteria is used. Using
a bespoke Matlab software has demonstrated differences in commercial software implementations.
As such, it is important to understand the response and limitations of the gamma index analysis
combined with the equipment in use. It was found for the same pass-rate criteria, different devices
and software combinations exhibit varying levels of agreement with each other. Therefore it is not
possible to recommend universal passing criteria for detector arrays or an optimal device
configuration; however this thesis has shown that a methodology incorporating Receiver Operator
Characteristic (ROC) analysis, in conjunction with a range of test plans with deliberately introduced
changes, could be used to derive optimal passing criteria in order to detect clinically relevant
deviations. Using a commercial detector array for a dosimetry audit of rotational radiotherapy is
suitable in place of standard systems of dosimetry.
11 FUTURE WORK

11.1 INTRODUCTION

This thesis has highlighted that detector arrays are suitable for verification of advanced IMRT given careful implementation. There are additional research topics that have arisen from this work that merit further investigations which could be performed. The following headings give descriptions for further work that should be investigated.

11.2 INVESTIGATION OF THE CHAPTER 9 METHODOLOGY IN VARIOUS COMMERCIAL DETECTOR ARRAY SYSTEMS

Extension of the methodology in Chapter 9 to other commercial detector array systems would ascertain if similar results could be achieved, independent of the measurement system. Additionally it would be of interest to compare other gamma index metrics such as the mean gamma in the different systems. The deliberately modified plans could be measured in the OCTAVIUS II and OCTAVIUS4D phantom to give 2D and 3D measurements which would give a direct comparison against the simulation results. This work would allow for more global recommendations on various devices and the appropriate acceptance criteria that could be considered, and may allow for the harmonisation of analysis techniques. This could be of benefit in dosimetry audits, particularly those that rely on individual centres to perform their own measurements of a standard treatment plan using local equipment [167].

11.3 FURTHER EVALUATION OF DETECTOR ARRAYS FOR 3D DOSE GUIDED RECONSTRUCTION IN PATIENT ANATOMY

The development of the SRS-1000 is a promising step in the detector array technology. With a spacing of 2.5mm in the central 5.5cm x 5.5cm area, this is similar to the typical voxel spacing in treatment planning systems providing better accuracy in the dose reconstruction as demonstrated in this study. The major limitation is the size of the array which would be limited to small treatment
regions. This thesis showed that sparse detector arrays may increase measurement uncertainty when used to reconstruct 3D dose distributions, however further studies with a number of different plans with different geometries are required for confirmation.

Additionally, algorithms have been developed to use the 3D measured data to back-project a dose distribution within patient anatomy. The accuracy of these algorithms and their comparison against gamma analysis has become a topic of interest. The measurements made in this thesis using the OCTAVIUS 4D phantom would be suitable for such a future comparison.

11.4 Correlation of Gamma Index Against Machine Specific Metrics
At the time of writing, a separate collaboration with Belfast Cancer Centre was underway to investigate the use and analysis of machine Dynalog file acquisition in an audit setting. Dynalog files are text files (on the Varian Clinac series) or binary files (on Varian TrueBeam models) which record the machine state every 20ms. Within these files is information about where the MLCs are positioned during those intervals, and the number of MUs recorded. The information is recorded against the corresponding expected values from the treatment plan. The information in the Dynalogs can then be interrogated to evaluate the accuracy of the delivery. This work has been linked to the national rotational radiotherapy audit and limited to Varian linear accelerators. As part of this, a standard 3DTPS RapidArc plan generated by the author at Royal Surrey County Hospital was delivered at different centres with Varian linac technology. During the audits, as well as measurements using the 2D-ARRAY 729, dynalog acquisition was switched on. The aim of this is to investigate the ability of different Varian linacs to deliver the same MLC leaf pattern. Work is being carried out to correlate the results from the 2D-ARRAY 729 with the dynalog analysis. The analysis of the dynalogs files is being led by Belfast and correlations will be attempted with the metrics investigated in Chapter 9.

11.5 Correlation of Gamma Index Against Treatment Plan Specific Metrics
Further work should also be carried out to assess whether the different gamma index metrics could be used to correlate against treatment plan specific parameters such as the number of MU or against metrics that score the complexity of a particular plan [159,160]. If there are statistically significant correlations between these parameters then it may be possible to develop a formalism where treatment plan specific metrics could be used to predict whether a certain plan is likely to fail.
an in-phantom verification measurement. This could be used to adjust the treatment plans appropriately and therefore allows for further streamlining in IMRT/VMAT QA by not having to perform a patient-specific measurement on every treatment plan.

11.6 CAN DETECTOR ARRAY DATA GIVE MEANINGFUL INFORMATION ABOUT THE ACCURACY OF LOW DOSE MODELLING?

Whilst setting a lower threshold for the gamma index calculation has been shown in Chapter 9 to be justified to avoid skewing the result, the disadvantage is that it may hide errors in the accuracy with which the low dose TPS modelling has been configured; particularly the leakage component of the MLC modelling. This part of the beam model will have a lower impact on a high dose PTV but is nonetheless a potentially unavoidable source of uncertainty. There may be more impact for OARs. Further work could be carried out using the database of measurements from the national Tomotherapy and VMAT audit and focussing on the accuracy of the predicted TPS dose in the low dose regions against the measured dose.

11.7 WHAT THRESHOLD VALUE SHOULD BE USED FOR ROC ANALYSIS TO DERIVE GAMMA INDEX ACCEPTANCE CRITERIA?

As shown in Chapter 8, it is possible to use ROC analysis to prospectively derive appropriate gamma index acceptance criteria through a simulation study. A key parameter in the ROC analysis was the setting of the threshold value for an out-of-tolerance ΔD%. In this work, this has been set at ±3% [28]. Work is needed to ascertain what level of dose difference is clinically acceptable for IMRT and VMAT treatments. This is a complex study that would require retrospective analysis of dosimetry data alongside patient related outcome data. A threshold value could then be used more accurately in a ROC analysis.

11.8 EMERGING APPROACHES IN ADVANCED IMRT VERIFICATION

There is currently growing interest in new approaches for IMRT verification, notably in vivo dosimetry using the EPID [168–171]. This technique involves measuring a transit dose by positioning the EPID device behind the patient on-treatment. The measured transit fluences are then back-projected onto an image of the patient anatomy using novel reconstruction algorithms to estimate the actual dose distribution within the patient. These techniques require evaluation and one method that would be interesting to investigate would be to use a detector array for the verification of the in vivo system. The deliberately modified plans developed in Chapter 4 could be used. An in vivo
measurement could be performed with a detector array / phantom combination, with the array simultaneously performing a measurement. The transit dose is reconstructed onto the phantom and this could potentially allow a direct comparison against the measured dose by the array. This would provide a novel methodology for verifying the accuracy of the in vivo system. Potential pitfalls will need to be carefully considered such as the inhomogeneities in the detector array system when measuring a transit dose. It must be remembered that the detector array system has been designed to measure dose in water and the calibration procedures are designed to cancel out the inhomogeneity effects. In this regard, 3D polymer gel dosimetry would be a solution as a true 3D dose distribution could be measured. As the whole dosimeter is tissue-equivalent there would be no issues with back-projecting the transit dose onto it for a like-for-like comparison against the measured distribution.

11.9 INVESTIGATION OF OTHER EVALUATION METRICS

Other metrics have been briefly discussed in Chapter 2, e.g. chi index, kappa index etc. There should be ongoing work to update the bespoke gamma index code appropriately to keep up with developments in the field; for example Sumida et al have recently proposed a novel radiobiological gamma index [172]. It would be interesting to use the measured data from Chapter 9 to perform a comparison between different metrics and whether the methodology described could be applied to them.

11.10 EVALUATION OF THE OCTAVIUS1500 DETECTOR ARRAY FOR 3D DOSE RECONSTRUCTION

The work carried out in Chapter 7 on 3D measurement guided dose reconstruction using the OCTAVIUS4D phantom, could be expanded to include the recently released (October 2014) OCTAVIUS1500 detector array. This work made use of the SRS1000 array which has a detector-to-detector spacing of 2.5mm and the 2D-ARRAY 729 which has a 10mm spacing. It was found that the 2D-ARRAY 729 caused artefacts in the 3D dose reconstruction due to the wide spacing. The OCTAVIUS1500 detector array has a spacing of 7mm and therefore it would be of interest to investigate whether it provides a reasonable 3D dose reconstruction, particularly as it covers a larger measurement area of 27cm x 27cm which could potentially offset the limited size of the SRS1000 array.
11.11 CAN THE BESPOKE MATLAB SOFTWARE BE USED TO DETERMINE OPTIMUM DETECTOR ARRAY CONFIGURATION?

The bespoke Matlab gamma index software described in Chapter 5 is configurable to the user’s needs with multiple options for changing different parameters. So far it is able to handle planar and 3D DICOM dose cube data. One interesting approach could be utilising the code in determining an optimum detector array configuration. The software would need to be altered so that it can effectively handle any type of detector array that should be possible to set by the user. Variables that would need to be included would be shape, size and detector spacing. The proposed methodology should be that gamma index calculations be first performed in an idealised situation using different plans with deliberate changes (as described in this thesis). Then calculations should be performed using practical configuration to determine a theoretical optimal design. Considerations would then need to be made on the practicalities and limitations of creating such a device, including cost, type of detector required, physical constraints and so forth. As such the scope of this type of work would be suitable for PhD thesis and could potentially be supported by an interested manufacturer.
REFERENCES


References


References


References


References


References


References


References


Appendix A  Basic Detector Array Commissioning Tests

A.1  Introduction

Basic commissioning tests of the detector arrays used throughout this thesis were performed on a Varian Clinac iX (Varian Medical Systems, Palo Alto, CA). The Clinac incorporates the Millennium 120 leaf MLCs, with the central 80 MLCs covering a 20x20cm each having a 0.5cm width at the isocentre; the remaining MLCs have 1cm width. All measurements were for 6 MV beam energy. The methodology was in keeping with previously published reports [60,66,74,88–94]. All systems were used according to individual manufacturer recommendations. For this section the methodology is described and results given for the 2D-ARRAY 729; this was the most heavily used detector array in this thesis and the configuration of it and the associated OCTAVIUS II phantom required the most stringent tests. Measurements for the other detector array systems agreed well with the published literature reports [60,66,74,88–94].

A.1.1 Effective point of measurement

The effective point of measurement (EPOM) was determined following the methodology of Poppe et al [89] and Van Esch et al [62]. For the 2D-ARRAY 729 The EPOM is specified in the manual as being 0.75cm from the surface. This places the point centrally between the parallel plate electrodes. Some groups in the literature report that the EPOM is 0.5cm from the surface, i.e. that it is at the top electrode [89]. In order to validate the EPOM for the 2D-ARRAY, an independent measurement was carried out. This was performed by placing various thicknesses of solid water on top of the 2D-ARRAY. There is a slab of 5mm PMMA on the top electrode. Assuming that the relative electron density of PMMA is 1.18, a water equivalent depth of 5.9cm was used. The settings used were 6MV, 100cm FSD, 10x10 cm field size, and 100 MU. Measurements were recorded for the central chamber. The readings were then normalised to the depth of dose maximum. A percentage depth dose (PDD) curve was then plotted using the effective water depth. The PDD measured by a diode in water was then plotted on the same graph. The shift in the two profiles was then calculated.

A.1.2 Dosimetric linearity

Dose linearity was determined between 5 – 2500 cGy. Dose rate linearity was checked for dose rates ranging from 100 – 600 MU/min. Output versus field size was checked for 2x2 – 25x25 field sizes. For field sizes of 5x5cm, 10x10, and 25x25cm, profiles in the 2D-ARRAY were compared with diode data measured in
Appendix A: Basic commissioning of detector arrays

a Scanditronix Wellhöfer water tank at the same depth. For all the tests, except EPOM, the setup was such that the effective depth of the 2D-ARRAY was at 5cm. The beam energy used throughout was 6MV (quality index, QI=0.670), focus-to-surface-distance (FSD) was 100cm, and 100 MU was used in all cases except the dose linearity check. For the linearity measurements, a 10x10cm field size was used.

The linearity of the 2D-ARRAY 729 with varying dose per pulse was also assessed. The dose per pulse was varied by changing the focus-to-surface distance (FSD). At each FSD, the field size at the surface was kept at a constant 10x10 in order to maintain the same scattering conditions. The inverse square law was used to set the collimators on the machine. The dose-per-pluse was previously measured using a Farmer chamber in solid water at 2cm depth, for a 10x10cm field size and 100cm FSD. The 2D-ARRAY 729 was setup in this condition for the measurements. At 6MV it was measured to be 0.033 cGy per pulse. Correction for different collection efficiency was calculated but was negligible. Table A.1 shows the FSD, field sizes that were used.

<table>
<thead>
<tr>
<th>FSD (cm)</th>
<th>Dose per pulse (cGy/pulse)</th>
<th>Machine field size setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>0.041</td>
<td>11.1 x 11.1</td>
</tr>
<tr>
<td>100</td>
<td>0.033</td>
<td>10.0 x 10.0</td>
</tr>
<tr>
<td>110</td>
<td>0.027</td>
<td>9.1 x 9.1</td>
</tr>
<tr>
<td>120</td>
<td>0.023</td>
<td>8.3 x 8.3</td>
</tr>
<tr>
<td>130</td>
<td>0.020</td>
<td>7.7 x 7.7</td>
</tr>
</tbody>
</table>

Measurements were performed using both the Farmer chamber at 5cm depth and 2D-ARRAY 729 at 5cm depth. Readings were normalised to the 100cm FSD measurement.

Furthermore, the detectors within the array have a relative calibration against the central detector and this was confirmed by setting an isocentric field with a 27x27cm field size to cover all the detectors within the array and assessing uniformity by looking at a profile through each line of detectors.

A.1.3 Field size comparisons

The 2D-ARRAY was compared with the RFA Linear Diode Array (LDA). The LDA is an array of 25 photon diodes, spaced 1cm centre-to-centre, providing similar measuring conditions to the 2D-ARRAY. Using the LDA and RFA water tank, profiles were taken for an open 20x20 cm 6MV field, a 20x20 cm field with 15° enhanced dynamic wedge (EDW), and a 20x20 cm field with 60° EDW. Measurements were performed
Appendix A: Basic commissioning of detector arrays

with 100 cm FSD and with the LDA at 5cm depth. A similar measuring scenario was recreated with the 2D-ARRAY and solid water. LDA and 2D-ARRAY 729 profiles were compared with the single diode full profile.

A.1.4 Comparisons of using the OCTAVIUS scan with 2D-ARRAY 729 in situ vs homogeneous scan.

Calculating on the scan of the OCTAVIUS with the 2D-ARRAY 729 in situ with an advanced calculation algorithm may result in perturbation of the predicted dose by the air filled ionization chambers, which may add to uncertainties, particularly when using the gamma index analysis [45]. Therefore, a dosimetric comparison was performed between using the OCTAVIUS scan with the 2D-ARRAY 729 in situ and a homogeneous insert.

Firstly, the directional response of the OCTAVIUS phantom was assessed by delivering a 10x10cm field in 15° gantry angle increments at 6MV with the phantom setup isocentrically. The dose to the central detector was recorded. To avoid irradiating through the couch, the sectors comprising the first 180° were measured with the OCTAVIUS phantom in the normal setup, and the remaining sectors were measured by inverting the phantom. The expected dose at the central detector was calculated in the Varian Eclipse™ v8.9 treatment planning system (Varian Medical Systems, Palo Alto, California) using the Analytical Anisotropic Algorithm (AAA) v8.9 algorithm [108] for both scans.

A 100 MU, 10x10 cm field was delivered isocentrically to the 2D-ARRAY 729 in the OCTAVIUS phantom from 24 gantry angles. These were delivered every 15°. Expected doses in the central ionization chamber of the 2D-ARRAY 729 were obtained from Eclipse. The plan was calculated on a CT scan of the OCTAVIUS phantom with the 2D-Array in situ in the coronal orientation. A comparison between expected central chamber dose and measured central chamber dose was performed. Calculations were performed using both the Pencil Beam Convolution (PBC) algorithm and the Analytical Anisotropic Algorithm (AAA). Irradiations were carried out with the 2D-ARRAY 729 setup in the OCTAVIUS phantom. For the posterior response, the 2D-Array and OCTAVIUS were turned upside-down and irradiated with anterior gantry angles to avoid measuring through the couch. The OCTAVIUS phantom has also been scanned with the homogeneous ionization chamber insert. Predicted doses were also calculated using this CT scan using the PBC and AAA.

Each gantry angle measurement was assessed separately. The total delivered dose was also compared with the total expected dose from all gantry angles.
Appendix A: Basic commissioning of detector arrays

A.2 Results and Discussion

A.2.1 Basic commissioning

A.2.2 Effective point of measurement

The percentage depth dose (PDD) measured by the 2D Array is shown plotted against the water tank beam data in Figure A.1. The difference between the 2D Array and the beam data was deduced to be a 2.5mm forward shift from the front electrode, i.e. an EPOM of 7.5mm from the surface of the 2D-ARRAY 729. Therefore the EPOM has been validated to be at the centre of the ionization chambers and is also in agreement with the published literature [62].

![Figure A.1. Comparison of the PDD for the 2D-ARRAY 729 and from the beam data](image)

A.2.3 Linearity of the 2D-ARRAY 729 with dose, dose-rate and dose-per-pulse

Dose linearity between 5 to 2500 cGy was excellent, as seen in in Figure A.2., and had a Pearson correlation coefficient, $r$, of 1.0. Dose rate linearity was found to be within ±0.2% between 100 – 600 MU/minute. Responses were normalised to 400 MU/min. Figure A.3. shows that there is no response relationship with varying dose rate. Figure A.4. shows the response as a function of dose per pulse for the 2D-ARRAY 729 and Farmer chamber.
Appendix A: Basic commissioning of detector arrays

Figure A.2. Linearity of the 2D-ARRAY 729 with monitor units

Figure A.3. Linearity of the 2D-ARRAY 729 with dose rate
Appendix A: Basic commissioning of detector arrays

A.2.4 Directional response in the OCTAVIUS scan with 2D-Array in situ vs homogeneous scan

The result of the directional response evaluation can be seen in Figure A.5. The graph shows the difference between the measured and expected dose in the central detector within the 2D-Array as a function of gantry angle. Table A.2. shows comparison of the total measured dose with predicted doses. The difference between the total dose given to the central detector and expected was 0.3% when the 2D-Array was scanned in the OCTAVIUS phantom and was -1.3% when a homogeneous insert was used. It can be seen that using a homogeneous scan results in a significant under-response when the beam incidence is lateral or entering the array through an oblique direction. This is due to the lack of modelling of the inhomogeneities caused by the vented ion chambers within the 2D-Array. When the scan of the 2D-Array was used, this improved the response. It is worth noting that this comparison is reported for the AAA algorithm. Calculating on the 2D-Array scan using the pencil beam convolution algorithm with heterogeneity correction yielded, as expected, a similar result to that seen with the homogeneous insert calculated using AAA.
Figure A.5. Angular response of the 2D-Array; comparison using predicted doses using PBC and AAA in the 2D-ARRAY 729 scan, and PBC and AAA calculations in the homogeneous scan.

Table A.2. Comparison of total dose at the centre chamber from all 10x10 100 MU beams delivered isocentrically from 24 gantry angles.

<table>
<thead>
<tr>
<th>2D Array measured</th>
<th>Predicted</th>
<th>2D Array scan</th>
<th>2D Array scan</th>
<th>Homogeneous</th>
<th>Homogeneous</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PBC</td>
<td>AAA</td>
<td>PBC</td>
<td>AAA</td>
<td></td>
</tr>
<tr>
<td>Dose at centre (cGy)</td>
<td>1478.3</td>
<td>1481.9</td>
<td>1467.3</td>
<td>1497.0</td>
<td>1500.6</td>
</tr>
<tr>
<td>% diff (measured – predicted)</td>
<td>-</td>
<td>-0.2</td>
<td>+0.7</td>
<td>-1.2</td>
<td>-1.5</td>
</tr>
</tbody>
</table>
A.2.5 Output vs field size response of the 2D-Array

The output versus field size response of the 2D array was output factors measured using the RK-chamber in a water tank. Both setups were to 5cm depth and 100cm FSD. Field sizes were delivered between $3 - 10$ cm$^2$ in 1 cm$^2$ increments and then 12, 14, 17, 20, and 25 cm$^2$. Figure A.6. shows the output factors for the 2D Array and Farmer chamber. Comparisons were also made with the output factors in the beam data charts. Excellent agreement can be seen between the 2D Array and chart data. The results show good agreement.

![Output vs field size comparison between the 2D-ARRAY 729 and beam data at 5cm depth in water, 100cm FSD.](image)

A.2.6 Comparison of the 2D-ARRAY 729 with diode profiles taken in the RFA Water tank

Profiles of field sizes 5 x 5cm, 10x10cm and 20 x 20 cm measured with the 2D-ARRAY 729 compared excellently with profiles measured using a diode in a water tank (the concordance correlation coefficient, $\rho_c$, was $> 0.999$ for all). Results for the 5x5cm field are shown in Figure A.7. The 10x10cm results are shown in Figure A.8. Results for the 20x20cm are shown in Figure A.9.
Appendix A: Basic commissioning of detector arrays

Figure A.7. (a) Comparison of 5cm field size measured by the 2D-ARRAY 729 and RFA water tank diode at 5 cm water-equivalent depth, 100 cm FSD.

Figure A.8. Comparison of 10cm field size measured by the 2D-ARRAY 729 and RFA water tank diode at 5 cm water-equivalent depth, 100 cm FSD.
Figure A.9. Comparison of 20cm field size measured by the 2D-ARRAY 729 and RFA water tank diode at 5 cm water-equivalent depth, 100 cm FSD.

A.2.7 Comparison of the 2D-ARRAY 729 with Linear Diode Array (LDA) profiles taken in the RFA plotting tank

The 2D Array was compared with a Linear Diode Array (LDA) in a water tank. The LDA is an array of 25 photon diodes, spaced 1cm centre-to-centre, providing similar measuring conditions to the 2D Array. Using the LDA and water tank, profiles were taken for an open 20x20 cm 6MV field, a 20x20 cm field with 15° enhanced dynamic wedge (EDW), and a 20x20 cm field with 60° EDW. Measurements were performed with 100 cm FSD and with the LDA at 5cm depth. A similar measuring scenario was recreated with the 2D Array and solid water. LDA and 2D Array profiles were compared with the single diode full profile as shown in Figures A.10 and A.11.
Figure A.10. Comparison between 2D-ARRAY 729, LDA, and single diode profile, 6 MV at dmax

Figure A.11. Comparison between 2D-ARRAY 729, LDA, and single diode profile, Wedge 15° Y1 direction, 6 MV at dmax. Comparison between 2D-ARRAY 729, LDA, and single diode profile, Wedge 60° Y1 direction, 6 MV at dmax.
A methodology for dosimetry audit of rotational radiotherapy using a commercial detector array.

Mohammad Hussein\textsuperscript{1,2,*}, Yatman Tsang\textsuperscript{3,4}, Russell Thomas\textsuperscript{5}, Clare Gouldstone\textsuperscript{5}, David Maughan\textsuperscript{5}, Julia AD Snaith\textsuperscript{5}, Steven C Bolton\textsuperscript{6,7}, Andrew Nisbet\textsuperscript{1,2}, Catharine H Clark\textsuperscript{1,4,5}.

\textsuperscript{1} Department of Medical Physics, Royal Surrey County Hospital, Guildford, UK, \textsuperscript{2} Centre for Nuclear and Radiation Physics, University of Surrey, UK, \textsuperscript{3} Mount Vernon Hospital, Middlesex, UK, \textsuperscript{4} NCRI Radiotherapy Trials QA Group, \textsuperscript{5} National Physical Laboratory, London, UK, \textsuperscript{6} The Christie Hospital, Manchester, UK, \textsuperscript{7} IPEM Audit group

Abstract

\textit{Purpose}: To develop a methodology for the use of a commercial detector array in dosimetry audits of rotational radiotherapy.

\textit{Materials and methods}: The methodology was developed as part of the development of a national audit of rotational radiotherapy. Ten cancer centres were asked to create a rotational radiotherapy treatment plan for a three-dimensional treatment-planning-system (3DTPS) test and audited. Phantom measurements using a commercial 2D ionization chamber (IC) array were compared with measurements using 0.125cm\textsuperscript{3} IC, Gafchromic film and alanine pellets in the same plane. Relative and absolute gamma index (γ) comparisons were made for Gafchromic film and 2D-Array planes respectively.

\textit{Results}: Comparisons between individual detectors within the 2D-Array against the corresponding IC and alanine measurement showed a statistically significant concordance correlation coefficient (both \(\rho>0.998\), \(p<0.001\)) with mean difference of \(-1.1\%\pm1.1\%\) and \(-0.8\%\pm1.1\%,\) respectively, in a high dose PTV. In the \(\gamma\) comparison between the 2D-Array and film it was that the 2D-Array was more likely to fail planes where there was a dose discrepancy due to the absolute analysis performed.

\textit{Conclusions}: It has been found that using a commercial detector array for a dosimetry audit of rotational radiotherapy is suitable in place of standard systems of dosimetry.

A critical evaluation of the PTW 2D-Array seven29 and Octavius II phantom for IMRT and VMAT verification.


Mohammad Hussein,\textsuperscript{1,2*} Elizabeth J. Adams,\textsuperscript{1} Thomas J. Jordan,\textsuperscript{1} Catharine H. Clark,\textsuperscript{1,3} and Andrew Nisbet\textsuperscript{1,2}
ABSTRACT
Quality assurance (QA) for intensity- and volumetric-modulated radiotherapy (IMRT and VMAT) has evolved substantially. In recent years, various commercial 2D and 3D ionization chamber or diode detector arrays have become available, allowing for absolute verification with near real time results, allowing for streamlined QA. However, detector arrays are limited by their resolution, giving rise to concerns about their sensitivity to errors. Understanding the limitations of these devices is therefore critical. In this study, the sensitivity and resolution of the PTW 2D-ARRAY seven29 and OCTAVIUS II phantom combination was comprehensively characterized for use in dynamic sliding window IMRT and RapidArc verification. Measurement comparisons were made between single acquisition and a multiple merged acquisition techniques to improve the effective resolution of the 2D-ARRAY, as well as comparisons against GAFCHROMIC EBT2 film and electronic portal imaging dosimetry (EPID). The sensitivity and resolution of the 2D-ARRAY was tested using two gantry angle 0° modulated test fields. Deliberate multileaf collimator (MLC) errors of 1, 2, and 5 mm and collimator rotation errors were inserted into IMRT and RapidArc plans for pelvis and head & neck sites, to test sensitivity to errors. The radiobiological impact of these errors was assessed to determine the gamma index passing criteria to be used with the 2D-ARRAY to detect clinically relevant errors. For gamma index distributions, it was found that the 2D-ARRAY in single acquisition mode was comparable to multiple acquisition modes, as well as film and EPID. It was found that the commonly used gamma index criteria of 3% dose difference or 3 mm distance to agreement may potentially mask clinically relevant errors. Gamma index criteria of 3%/2 mm with a passing threshold of 98%, or 2%/2 mm with a passing threshold of 95%, were found to be more sensitive. We suggest that the gamma index passing thresholds may be used for guidance, but also should be combined with a visual inspection of the gamma index distribution and calculation of the dose difference to assess whether there may be a clinical impact in failed regions.

PACS numbers: 87.55.Qr, 87.56.Fc
Key words: IMRT, VMAT, QA, detector arrays

A comparison of the gamma index analysis in various commercial IMRT/VMAT QA systems


Mohammad Hussein1,2, Pejman Rowshanfarzad2, Martin A Ebert3,4, Andrew Nisbet1,2, Catharine H Clark1,5

1 Department of Medical Physics, Royal Surrey County Hospital, Guildford, UK, 2 Centre for Nuclear and Radiation Physics, University of Surrey, UK, 3 School of Physics, University of Western Australia, Crawley, Western Australia, Australia, 4 Department of Radiation Oncology, Sir Charles Gairdner Hospital, Nedlands, Western Australia, Australia, 5 National Physical Laboratory, London, UK

Abstract

Purpose

The purpose of this study was to investigate the variability of the global gamma index (γ) analysis in various commercial IMRT/VMAT QA systems and to assess the impact of measurement with low resolution detector arrays on γ.

Materials
Five commercial QA systems (PTW 2D-Array, Scandidos Delta4, SunNuclear ArcCHECK, Varian EPID, and Gafchromic EBT2 film) were investigated. The response of γ analysis to deliberately introduced errors in pelvis and head & neck IMRT and RapidArc™ plans was evaluated in each system. A theoretical γ was calculated in each commercial QA system software (PTW 2D-Array, Scandidos Delta4, SunNuclear ArcCHECK, Varian EPID, and Gafchromic EBT2 film), using treatment planning system resolution virtual measurements and compared to an independent calculation. Error-induced plans were measured on a linear accelerator and were evaluated against the error-free dose distribution calculated using Varian Eclipse™ in the relevant phantom CT scan. In all cases, global γ was used with a 20% threshold relative to a point selected in a high dose and low gradient region. The γ based on measurement was compared against the theoretical to evaluate the response of each system.

Results

There was statistically good agreement between the predicted γ based on the virtual measurements from each software (concordance correlation coefficient, ρc>0.92) relative to the independent prediction in all cases. For the actual measured data, the agreement with the predicted γ reduces with tightening passing criteria and the variability between the different systems increases. This indicates that the detector array configuration and resolution have greater impact on the experimental calculation of γ due to under-sampling of the dose distribution, blurring effects, noise, or a combination.

Conclusions

It is important to understand the response and limitations of the gamma index analysis combined with the equipment in use. For the same pass-rate criteria, different devices and software combinations exhibit varying levels of agreement with the predicted γ analysis.

B.1 Full publication history throughout duration of PhD

B.1.1 Peer reviewed publications


B.1.2 Published conference abstracts


Hussein M, Clark C H & Nisbet A. Detection of errors in various commercial IMRT/VMAT QA systems. IMRT Verification: making the most of it, IPEM/RTSIG Meeting. (Invited oral presentation)


Appendix C  Gamma Index Software
Matlab Code

% --- Routine to import data set1 (similar code for data set 2).
% --- Select Data set 1
function select_file_Callback(hObject, eventdata, handles)
clc;
file2=handles.file2;
if file2~=0
    [direct,file]=uigetfile({'*.dcm','DICOM files (*.dcm)';'*.xls','Microsoft Excel 97-2003 (*.xls)';'*.mcc','PTW 2D-Array measurement (*.mcc)';...
    '*.x','All files (*.*)'},'Select a file to load into dataset 1', file2);
else
    [direct,file]=uigetfile({'*.dcm','DICOM files (*.dcm)';'*.xls','Microsoft Excel 97-2003 (*.xls)';'*.mcc','PTW 2D-Array measurement (*.mcc)';...
    '*.x','All files (*.*)'},'Select a file to load into dataset 1');
end
path1=strcat(file,direct);
 [~, ~, ext1] = fileparts(path1);
handles.ext1=ext1;
direct1=direct;
handles.file1=file;
if strcmp(ext1,'.dcm')==1,
    dicom_file=dicomread(path1);
    infol=dicominfo(path1);
    handles.infol=inf1;
    handles.spacing_1=inf1.PixelSpacing(1);
    if strcmp(infol.Manufacturer, 'Math Resolutions, LLC')==1,
        interpfaor=1;
    else
        interpfactor=0;
    end
end
% check if dicom is 2D plane or 3D cube
% num_dims: 2(=2D), 4(=3D)
num_dims1=ndims(dicom_file);

handles.num_dims1=num_dims1;

% if dicom_file is a cube
if num_dims1==4,
    info2=handles.info2;
cube=squeeze(dicom_file);
cube=double(cube);
cube=cube*info1.DoseGridScaling;
cube=interp3(cube, interpfactor);

check_info1=exist('info1.Manufacturer');
if check_info1==1
    if strcmp(info2.Manufacturer, 'Math Resolutions, LLC')==1,
        cube2=handles.cube2;
        [x, y, z]=size(cube2);
        cube_expand=zeros(x, y, z);
        cube_expand(3:end-1,2:end-1,2:end-1)=cube;
        cube=cube_expand;
    else
    end
else
end
end

handles.file=file;
handles.cube=cube;
handles.cube_plot=cube;
handles.data1_orientation_val=1;

max_dose=max(cube(:));

set(handles.data1_orientation,'enable','on');
set(handles.slider1,'enable','on');

sizeImg = size(cube);
minSlice = 1;
maxSlice = sizeImg(3);
startSlice = round((maxSlice-minSlice)/2);

handles.slice = startSlice;

sliderStep = [1, 1] / (maxSlice - minSlice);

set(handles.slider1,'value',startSlice); %
set(handles.slider1,'max',maxSlice); %
set(handles.slider1,'min',minSlice);
set(handles.slider1, 'SliderStep', sliderStep);
set(handles.slicedepth_max,'value',maxSlice); %
set(handles.slicedepth_min,'value',minSlice); %

plane=cube(:,:,startSlice);
handles.plane=plane;

else

plane=double(dicom_file);
plane=plane*info1.DoseGridScaling/10;
max_dose=max(plane(:));
handles.plane=plane;

end

axes(handles.axes1);
imagesc(plane);
colormap(jet)
caxis([0 max_dose])
freezeColors
axis equal
axis tight

[sizex,sizey]=size(plane);
sizeplane = size(plane);
minplane = 1;
maxplane = sizeplane(1);
startplane = round((maxplane-minplane)/2);
planeStep = [1, 1] / (maxplane - minplane);

set(handles.profile_scroll,'value',startplane); %
set(handles.profile_scroll,'max',maxplane); %
set(handles.profile_scroll,'min',minplane);
set(handles.profile_scroll, 'SliderStep', planeStep);

axes(handles.profile_axes)
plot(plane(round(sizex/2),:))
hold

message=strcat(direct,' loaded.');
set(handles.Directory,'String',message);
clc

elseif strcmp(ext1,'.xls')==1

%%
% Import the data
handles.num_dims1=2;
if strcmp(direct,'Matlab OCTAVIUS 729 coronal.xls')==1

    [~, ~, raw] = xlsread(path1,'Normal','B1:CZ105');
elseif strcmp(direct,'Matlab OCTAVIUS SRS1000 coronal.xls')==1

    [~, ~, raw] = xlsread(path1,'TPS','B1:AR41');
else

    [~, ~, raw] = xlsread(path1,'TPS','B1:AB27');

end
end

% Create output variable
xls_data = cell2mat(raw);
% Clear temporary variables
clearvars raw;

xls_data=flipud(xls_data);
%xls_data=interp2(xls_data,3,'spline');

max_dose=max(xls_data(:));

handles.xls_data=xls_data;

plane=xls_data;
handles.plane=plane;
[sizex,~]=size(plane);

sizeplane = size(plane);
minplane = 1;
maxplane = sizeplane(1);
startplane = round((maxplane-minplane)/2);
planeStep = [1, 1] / (maxplane - minplane);

set(handles.profile_scroll,'value',startplane); %
set(handles.profile_scroll,'max',maxplane); %
set(handles.profile_scroll,'min',minplane);
set(handles.profile_scroll, 'SliderStep', planeStep);

axes(handles.profile_axes)
plot(plane(round(sizex/2),:))
hold
clc

axes(handles.axes1);
imagesc(xls_data);
colormap(jet)
caxis([0 max_dose])
freezeColors

axis equal
axis tight

message=strcat(direct,' loaded. ');
set(handles.Directory,'String','message');

end
guidata(hObject, handles);

% --- Executes on selection change in data1_orientation.
% % Pop-up menu to select axial, coronal or sagittal plane
% function data1_orientation_Callback(hObject, eventdata, handles)

ext1=handles.ext1;
ext2=handles.ext2;

if strcmp(ext1,'.dcm')==1,
data1_orientation_val = get(hObject, 'Value');
handles.data1_orientation_val = data1_orientation_val;

set(handles.data2_orientation, 'Value', data1_orientation_val);
data2_orientation_Callback(hObject, eventdata, handles)
cube = handles.cube;

if data1_orientation_val == 1
cube_plot = cube;
elseif data1_orientation_val == 2
   cube_plot = permute(cube, [3 2 1]);
elseif data1_orientation_val == 3
   cube_plot = permute(cube, [3 1 2]);
end

max_dose = max(cube(:));
handles.cube_plot = cube_plot;

sizeImg = size(cube_plot);
minSlice = 1;
maxSlice = sizeImg(3);
startSlice = round((maxSlice - minSlice)/2);

handles.slice = startSlice;
sliderStep = [1, 1] / (maxSlice - minSlice);

set(handles.slider1, 'value', startSlice); %
set(handles.slider1, 'max', maxSlice); %
set(handles.slider1, 'min', minSlice);
set(handles.slider1, 'SliderStep', sliderStep);

set(handles.slicethickness_max, 'string', maxSlice); %
set(handles.slicethickness_min, 'string', minSlice); %
set(handles.slicethickness_current, 'string', startSlice); %

if data1_orientation_val == 1
   axes(handles.axes1)
   imagesc(cube_plot(:,:,startSlice));
   colormap(jet)
   caxis([0 max_dose])
   freezeColors
   axis equal
   axis tight
else
   axes(handles.axes1)
   imagesc(cube_plot(:,:,startSlice));
   colormap(jet)
   caxis([0 max_dose])
   freezeColors
   axis equal
   axis tight
end
plane=cube_plot(:,:,startSlice);
handles.plane=plane;

[sizex,sizey]=size(plane);
sizeplane = size(plane);
minplane = 1;
maxplane = sizeplane(1);
startplane = round((maxplane-minplane)/2);
planeStep = [1, 1] / (maxplane - minplane);

set(handles.profile_scroll,'value',startplane); %
set(handles.profile_scroll,'max',maxplane); %
set(handles.profile_scroll,'min',minplane);
set(handles.profile_scroll, 'SliderStep', planeStep);

axes(handles.profile_axes)
if strcmp(ext2,'.dcm')==1,
  hold
  plot(plane(round(sizex/2),:))
  hold
  clc
else
  plot(plane(round(sizex/2),:))
  hold
  clc
end
elseif strcmp(ext1,'.xls') ==1
  %
  % enter code here
  %
end

guidata(hObject, handles);

% --- Executes during object creation, after setting all properties. 
function data1_orientation_CreateFcn(hObject, eventdata, handles)
% 2D Gamma index calculation routine
% --- Executes on button press in calc_gamma. 
function calc_gamma_Callback(hObject, eventdata, handles)
tic;
local_global=get(handles.gamma_type,'Value');
interp_type=get(handles.interpolation_type, 'Value');
calc_2D_3D=get(handles.calc_2D_3D, 'Value');
search_dist=get(handles.search_dist,'String');
data1_orientation_val=get(handles.data1_orientation,'Value');
data2_orientation_val=get(handles.data2_orientation,'Value');
h=waitbar(0, 'Calculating....');
ext1=handles.ext1;
ext2=handles.ext2;
set(handles.display_map_result,'enable','on');
doseCriterion=get(handles.dose_diff,'String');
doseCriterion=str2double(doseCriterion)/100;
distanceCriterion=get(handles.DTA,'String');
distanceCriterion=str2double(distanceCriterion);
normDose=get(handles.normalisation,'String');
normDose=str2double(normDose);

% interpolation factor used for map 1
interpfactor=get(handles.interpfactor,'String');
interpfactor=str2double(interpfactor);

% interpolation factor used for map 2
interpfactor2=get(handles.interpfactor2,'String');
interpfactor2=str2double(interpfactor2);

%Interpolate maps
if interp_type==1,
    interp_text = 'spline';
elseif interp_type==2,
    interp_text = 'linear';
elseif interp_type==3,
    interp_text = 'cubic';
elseif interp_type==4,
    interp_text = 'nearest';
end

search_dist=search_dist*2^interpfactor;
if strcmp(ext1,'.dcm')==1,
    info1=handles.info1;
    num_dims1=handles.num_dims1;
    spacing_x_1=info1.PixelSpacing(1)/2^interpfactor;
    spacing_y_1=info1.PixelSpacing(2)/2^interpfactor;
    % check if file is cubic or 2D plane
    if num_dims1==4,
        slice=handles.slice;
Appendix C: Gamma index software matlab code extract

cube1=handles.cube;

[sizex_orig, sizey_orig, sizez_orig]=size(cube1);

if datal_orientation_val == 1 % Axial
    spacing_x_1=info1.PixelSpacing(1)/2^interpfactor;
    spacing_y_1=info1.PixelSpacing(2)/2^interpfactor;
    slice_thickness=info1.GridFrameOffsetVector(2)-
    info1.GridFrameOffsetVector(1)/2^interpfactor;
elseif datal_orientation_val == 2 % Coronal
    cube1=permute(cube1,[3 2 1]);
    spacing_y_1=info1.PixelSpacing(2)/2^interpfactor;
    slice_thickness=info1.PixelSpacing(1)/2^interpfactor;
    spacing_x_1=(info1.GridFrameOffsetVector(2)-
    info1.GridFrameOffsetVector(1))/2^interpfactor;
elseif datal_orientation_val == 3
    cube1=permute(cube1,[3 1 2]); % Sagittal
    slice_thickness=info1.PixelSpacing(2)/2^interpfactor;
    spacing_y_1=info1.PixelSpacing(2)/2^interpfactor;
    spacing_x_1=(info1.GridFrameOffsetVector(2)-
    info1.GridFrameOffsetVector(1))/2^interpfactor;
end

plane_orig=cube1(:,:,slice);

slice_interp=slice*2^interpfactor-(2^interpfactor-1);

plane=cube1(:,:,slice_interp);

else
    plane=handles.plane;
    plane_orig=plane;
    plane=interp2(plane, interpfactor , interp_text);
    [sizex, sizey]=size(plane);
end

elseif strcmp(ext1,'.xls')==1
    direct=handles.direct1;
    plane=handles.xls_data;
    num_dims1=2;
    plane_orig=plane;
    plane=interp2(plane, interpfactor, interp_text);

end
if strcmp(direct,'Matlab OCTAVIUS 729 coronal.xls')==1 ||
strcmp(direct,'Matlab OCTAVIUS SRS1000 coronal.xls')==1,
    spacing_x_1=2.5/2^interpfactor;
    spacing_y_1=2.5/2^interpfactor;
else
    spacing_x_1=10/2^interpfactor;
    spacing_y_1=10/2^interpfactor;
end
end

if strcmp(ext2,'.mcc')==1
    plane2=handles.array_plot;
    plane2_orig=plane2;
    plane2=interp2(plane2,interpfactor2,interp_text);
    [sizex_2,sizey_2]=size(plane2);
    spacing_x_2=10/2^interpfactor2;
    spacing_y_2=10/2^interpfactor2;
    cube_coronal=permute(cube1,[3 2 1]);
    waitbar(0,h,'Done processing....');
    [sizex, sizey, sizez]=size(cube_coronal);
    eval_slice=round(sizez/2-2^interpfactor);
    plane=cube_coronal(:,:,eval_slice);
    spacing_x_1=(info1.GridFrameOffsetVector(2)-
    info1.GridFrameOffsetVector(1))/2^interpfactor;
    spacing_y_1=info1.PixelSpacing(2) / 2^interpfactor;
    slice_thickness=info1.PixelSpacing(1) / 2^interpfactor;
    num_dims2=2;

    search_dist_z = round(distanceCriterion * 2 * 2^interpfactor);
    %diff_threshold=NaN(sizex_2, sizey_2);
    %lowercut=0.95;
    %uppercut=1;
    %diff_threshold(plane2)>=lowercut*max(plane2(:)) &
    plane2<=uppercut*max(plane2(:)))=1;
    %plane2_threshold=plane2.*diff_threshold;
    %normDose=nanmean(plane2_threshold(:));

    search_dist_x=round((sqrt(sizex_2^2+sizey_2^2)*2^interpfactor));
    search_dist_y=round((sqrt(sizex_2^2+sizey_2^2)*2^interpfactor));
    offset_x=(sizex_2*(abs(spacing_x_2/spacing_x_1))-((spacing_x_2/spacing_x_1)-1)-sizex;
    offset_y=(sizey_2*(abs(spacing_y_2/spacing_y_1))-((spacing_y_2/spacing_y_1)-1)-sizey;
elseif strcmp(ext2,'.dcm')==1
    info2=handles.info2;
    num_dims2=handles.num_dims2;
    if num_dims2==4,
        cube2=handles.cube2;
        [sizex_2, sizey_2, sizez_2]=size(cube2);
        spacing_x_2=info2.PixelSpacing(1)/2^interpfactor2;
        spacing_y_2=info2.PixelSpacing(2)/2^interpfactor2;
        offset_x=sizex_2-sizex_orig;
        offset_y=sizey_2-sizey_orig;
        if offset_x > 0 || offset_y > 0
            cube2=cube2(1+offset_x/2^interpfactor2:sizex_2,
                    1+offset_y/2^interpfactor2:sizey_2, 1:sizez_2);
        elseif offset_x < 0 && offset_y < 0
            cube2_expand=zeros(sizex_orig, sizey_orig, sizez_orig);
            cube2_expand(1:sizex_2, 1:sizey_2, :)=cube2;
            cube2=cube2_expand;
        else
        end
        cube2=interp3(cube2, interpfactor2, interp_text);
    if data2_orientation_val ==1 % Axial
        spacing_x_2=info2.PixelSpacing(1)/2^interpfactor2;
        spacing_y_2=info2.PixelSpacing(2)/2^interpfactor2;
    elseif data2_orientation_val == 2 % Coronal
        cube2=permute(cube2,[3 2 1]);
        spacing_y_2=info2.PixelSpacing(2)/2^interpfactor2;
        spacing_x_2=(info2.GridFrameOffsetVector(2)-
                      info2.GridFrameOffsetVector(1))/2^interpfactor2;
    elseif data2_orientation_val == 3
        cube2=permute(cube2,[3 1 2]); % Sagittal
        spacing_y_2=info2.PixelSpacing(1)/2^interpfactor2;
        spacing_x_2=(info2.GridFrameOffsetVector(2)-
                      info2.GridFrameOffsetVector(1))/2^interpfactor2;
    end
    slice=handles.slice;
    [sizex_2, sizey_2, sizez_2]=size(cube2);
    plane2=cube2(:,:,slice);
    handles.current_slice2=slice;
else
Appendix C: Gamma index software matlab code extract

```matlab
plane2=handles.plane2;
plane2_orig=plane2;
plane2=interp2(plane2,interpfactor2,interp_text);
[sizex_2, sizey_2]=size(plane2);

end

search_dist_z = round(distanceCriterion * 2 * 2^interpfactor);
search_dist_x=round((sqrt(sizex_2^2+sizey_2^2)*2^interpfactor));
search_dist_y=round((sqrt(sizex_2^2+sizey_2^2)*2^interpfactor));

elseif strcmp(ext2,'.xls')==1

direct=handles.direct1;
num_dims2=2;
plane2=handles.xls_data2;
plane2_orig=plane2;
plane2=interp2(plane2,interpfactor2,interp_text);
[sizex_2,sizey_2]=size(plane2);

if strcmp(direct,'Matlab OCTAVIUS 729 coronal.xls')==1 ||
    strcmp(direct,'Matlab OCTAVIUS SRS1000 coronal.xls')==1,
    spacing_x_2=2.5/2^interpfactor2;
    spacing_y_2=2.5/2^interpfactor2;
else
    spacing_x_2=10/2^interpfactor2;
    spacing_y_2=10/2^interpfactor2;
end

offset_x=(sizex_2*(abs(spacing_x_2/spacing_x_1))-
    ((spacing_x_2/spacing_x_1)-1)-sizex;
offset_y=(sizey_2*(abs(spacing_y_2/spacing_y_1))-
    ((spacing_y_2/spacing_y_1)-1)-sizey;

end

% Insert calculation code here -------------------------------------

if num_dims1==4 && num_dims2==4 && calc_2D_3D==1,
    eval_planel = zeros(sizex + search_dist_x * 2, sizey + search_dist_y * 2);
eval_planel(:, :) = -10^8;
eval_planel(search_dist_x + 1:search_dist_x + sizex,search_dist_y +
1:search_dist_y + sizey) = plane;
gValue=zeros(search_dist_x * 2 + 1, search_dist_y * 2 + 1);
gValue(:, :)=100;
gamma=zeros(sizex_2, sizey_2);
distdiff=zeros(search_dist_x * 2 + 1, search_dist_y * 2 + 1);
dose_diff=zeros(sizex_2, sizey_2);
for i=1:search_dist_x * 2 + 1
    for j=1:search_dist_y * 2 + 1
```
distdiff(i,j)=sqrt(sqrt(((i-search_dist_x-1)*spacing_x_1)^2+((j-search_dist_y-1)*spacing_y_1)^2)^2);
end
end

for x=1:sizex_2
    waitbar(x/sizex_2,h);
    if getappdata(h,'canceling')
        break
    end
    for y=1:sizey_2
        testdose=eval_plane1(1+(x-1)*spacing_x_2/spacing_x_1:(x-1)*spacing_x_2/spacing_x_1+search_dist_x*2+1,...
            1+(y-1)*spacing_y_2/spacing_y_1:(y-1)*spacing_y_2/spacing_y_1+search_dist_y*2+1);
        if local_global==1,
            test_dosediff=(plane2(x,y)-testdose)/normDose;
        else
            test_dosediff=(plane2(x,y)-testdose)./testdose;
        end
        gValue=sqrt((test_dosediff./doseCriterion).^2 + (distdiff ./ (distanceCriterion)).^2);
        gamma(x,y)=min(gValue(:));
    end
end

elseif num_dims1==4 && num_dims2==4 && calc_2D_3D==2,

    search_dist_z = 2;
    %search_dist_x=round((sqrt(sizex_2^2+sizey_2^2)*2^interpfactor));
    %search_dist_y=round((sqrt(sizex_2^2+sizey_2^2)*2^interpfactor));
    search_dist_x=6; %
    search_dist_y=6;

    eval_cube1=zeros(sizex + search_dist_x*2, sizey + search_dist_y*2, search_dist_z*2+1);
    eval_cube1(:,:, :)=-10^8;
    eval_cube1(search_dist_x + 1 : search_dist_x + sizex,...
        search_dist_y + 1 : search_dist_y + sizey,:) = ...
        cube1(:,:,slice_interp - search_dist_z : search_dist_z + slice_interp);
gValue=zeros(search_dist_x * 2 + 1, search_dist_y * 2 + 1, search_dist_z * 2 + 1);
gValue(:,:,1)=100;
distdiff=zeros(search_dist_x * 2 + 1, search_dist_y * 2 + 1, search_dist_z * 2 + 1);
gamma=zeros(sizex_2, sizey_2);

for k=1:search_dist_z * 2 + 1
    for i=1:search_dist_x * 2 + 1
        for j=1:search_dist_y * 2 + 1
            distdiff(i,j,k)=sqrt(sqrt(((i-search_dist_x-1)*spacing_x_1)^2+(((j-search_dist_y-1)*spacing_y_1)^2)^2+(k-search_dist_z-1)*slice_thickness)^2);
        end
    end
end

for x=1:sizex_2
    waitbar(x/sizex_2,h);
    if getappdata(h,'canceling')
        break
    end
    for y=1:sizey_2
        testdose=eval_cube1(1+(x-1)*spacing_x_2/spacing_x_1:(x-1)*spacing_x_2/spacing_x_1+search_dist_x*2+1,...
                       1+(y-1)*spacing_y_2/spacing_y_1:(y-1)*spacing_y_2/spacing_y_1+search_dist_y*2+1,:);
        if local_global==1,
            test_dosediff=(cube2(x,y,slice)-testdose)/normDose;
        else
            test_dosediff=(cube2(x,y,slice)-testdose)./testdose;
        end
        gValue=sqrt((test_dosediff./doseCriterion).^2 + (distdiff./distanceCriterion).^2);
        gamma(x,y)=min(gValue(:));
    end
end
elseif num_dims1==4 && num_dims2==2 && calc_2D_3D==2,

    gValue=zeros(search_dist_x * 2 + 1, search_dist_y * 2 + 1, search_dist_z * 2 + 1);
    gValue(:,:,:)=100;
    distdiff=zeros(search_dist_x * 2 + 1, search_dist_y * 2 + 1, search_dist_z * 2 + 1);
    gamma=zeros(sizex_2, sizey_2);
    dose_diff=zeros(sizex_2, sizey_2);

    if strcmp(ext2,'.mcc')==1, % perform 3D gamma analysis on 2D array data
        % provided a Dose Cube is imported
    for k=1:search_dist_z * 2 + 1
        for i=1:search_dist_x * 2 + 1
            for j=1:search_dist_y * 2 + 1
                distdiff(i,j,k)=sqrt(sqrt(((i-search_dist_x-1)*spacing_x_1)^2+((j-search_dist_y-1)*spacing_y_1)^2)^2+((k-search_dist_z-1)*slice_thickness)^2);
            end
        end
    end

    for x=1:sizex_2
        waitbar(x/sizex_2,h);
        if getappdata(h,'canceling')
            break
        end
        for y=1:sizey_2
            testdose =
                cube_coronal(5*(2^interpfactor)*x+(10*2^interpfactor)-(2^interpfactor-1)-
                search_dist:5*(2^interpfactor)*x+(10*2^interpfactor)-(2^interpfactor-1)+search_dist,...
                4*(2^interpfactor)*y+(9*2^interpfactor)-search_dist:
                4*(2^interpfactor)*y+(9*2^interpfactor)+search_dist,...
                eval_slice-search_dist_z: eval_slice+search_dist_z);
            test_dosediff=(plane2(x,y)-testdose)/normDose;
            gValue=sqrt(((test_dosediff./doseCriterion).^2 + (distdiff ./ (distanceCriterion)).^2));
            gamma(x,y)=min(gValue(:));

            % calculate TPS dose over the chamber area of 5mm x 5mm
            TPS_chamber=testdose(round(search_dist+1-
                spacing_x_2/spacing_x_1/4) : round(search_dist+1+spacing_x_2/spacing_x_1/4),...
                round(search_dist+1-spacing_y_2/spacing_y_1/4) : round(search_dist+1+spacing_y_2/spacing_y_1/4),...
round(search_dist_z + 1 - spacing_x_2 / slice_thickness / 4) :
round(search_dist_z + 1 + spacing_x_2 / slice_thickness / 4));

dose_diff(x,y) = ((plane2(x,y) -
mean(TPS_chamber(:))) / mean(TPS_chamber(:))) * 100;
end
end

gammamap(gamma)
diff_threshold = NaN(sizex_2, sizey_2);
lowercut = 0.95;
uppercut = 1;
diff_threshold(plane2 >= lowercut * max(plane2(:)) &
plane2 <= uppercut * max(plane2(:))) = 1;

plane2_threshold = plane2 .* diff_threshold;
dose_diff_threshold = dose_diff .* diff_threshold;
ICRU_mean = nanmean(dose_diff_threshold(:));
ICRU_median = nanmedian(dose_diff_threshold(:));
ICRU_SD = nanstd(dose_diff_threshold(:));
ICRU_count = nansum(dose_diff_threshold(:) >= -1000);

else

1;
end

%######################## 2D vs 2D gamma analysis ########################
elseif num_dims1 == 2 && num_dims2 == 2,

    calc_type = 2;

eval_plane1 = zeros(sizex + search_dist * 2, sizey + search_dist * 2);
eval_plane1(:, :) = -10^8;
eval_plane1(search_dist + 1:search_dist + sizex, search_dist + 1:search_dist + sizey) = plane;

gValue = zeros(search_dist * 2 + 1, search_dist * 2 + 1);
gValue(:, :) = 100;
gamma = zeros(sizex_2, sizey_2);
distdiff = zeros(search_dist * 2 + 1, search_dist * 2 + 1);
dose_diff = zeros(sizex_2, sizey_2);

DTA_map = zeros(sizex_2, sizey_2);
DTA_map(:, :) = 2 * distanceCriterion;
dose_diff_map = zeros(sizex_2, sizey_2);
CI_map = zeros(sizex_2, sizey_2);
kappa_map = zeros(sizex_2, sizey_2);

if strcmp(ext2, '.dcm') == 1 || strcmp(ext2, '.xls') == 1,
if calc_type == 1,
    % perform calculation using search of whole evaluated distribution

    for x = 1:sizex_2
waitbar(x/sizex_2,h);
if getappdata(h,'canceling')
    break
end
for y=1:sizey_2
    for i=1:sizex
        for j=1:sizey
            test_dosediff=(plane2(x,y)-plane)/normDose;
            distdiff=sqrt(((x-1)*spacing_x_2-((i-1)*spacing_x_1))^2+((y-1)*spacing_y_2-((j-1)*spacing_y_1))^2);
            gValue(i,j)=sqrt((test_dosediff(i,j)./doseCriterion).^2 + (distdiff ./ (distanceCriterion)).^2);
        end
        end
        gamma(x,y)=min(gValue(:));
    end
end
elseif calc_type==2,
    % perform calculation using limited search distance specified in mm
    for i=1:search_dist * 2 + 1
        for j=1:search_dist * 2 + 1
            distdiff(i,j)=sqrt(sqrt(((i-search_dist-1)*spacing_x_1))^2+((j-search_dist-1)*spacing_y_1))^2+((i-search_dist-1)*spacing_x_1))^2);  
        end
    end
    for x=1:sizex_2
        waitbar(x/sizex_2,h);
        if getappdata(h,'canceling')
            break
        end
        for y=1:sizey_2
            testdose=eval_plane1(1+(x-1)*spacing_x_2/spacing_x_1:(x-1)*spacing_x_2/spacing_x_1+search_dist*2+1,...

```
1+(y-1)*spacing_y_2/spacing_y_1:(y-1)*spacing_y_2/spacing_y_1+search_dist*2+1);

if local_global==1,
    test_dosediff=(plane2(x,y)-testdose)/normDose;
else
    test_dosediff=(plane2(x,y)-testdose)./testdose;
end

gValue=sqrt((test_dosediff./doseCriterion).^2 + (distdiff ./ (distanceCriterion)).^2);

gamma(x,y)=min(gValue(:));

end
end
end
end

%------------------------------------------------------------------------------------------------------------------------
time_taken=num2str(toc,'%.2f');
diff_threshold=NaN(sizex_2, sizey_2);
diff_threshold_plot=zeros(sizex_2, sizey_2);

lowercut=get(handles.lower_cut,'String');
lowercut=str2double(lowercut)/100;

uppercut=get(handles.upper_cut,'String');
uppercut=str2double(uppercut)/100;

diff_threshold(plane2>=lowercut*max(plane2(:))&
plane2<=uppercut*max(plane2(:)))=1;
diff_threshold_plot(plane2>=lowercut*max(plane2(:))&
plane2<=uppercut*max(plane2(:)))=1;

plane2_threshold=plane2.*diff_threshold;

if interpfactor2==0,
    plane_threshold=plane_orig.*diff_threshold;
elseif interpfactor2==interpfactor,
    plane_threshold=plane.*diff_threshold;
end

gamma_thresh=gamma.*diff_threshold;

handles.gamma=gamma;
handles.gamma_plane=gamma;
handles.gamma_thresh = gamma_thresh;

% Compute cumulative histogram

gammavert = gamma_thresh(:);

k = 0;
if max(gammavert) > 5,
    for j = 0:0.05:5
        k = k + 1;
        cumul(k) = sum(gammavert >= j) / sum(gammavert >= 0) * 100; %#ok<AGROW>
        gammaval(k) = j; %#ok<AGROW>
    end
else
    for j = 0:0.05:max(gammavert)
        k = k + 1;
        cumul(k) = sum(gammavert >= j) / sum(gammavert >= 0) * 100; %#ok<AGROW>
        gammaval(k) = j; %#ok<AGROW>
    end
end

gamma_cumul = cumul';
gammavalt = gammaval';

sort_gamma_cumul = sort(gamma_cumul);
[x_cumul, ~] = size(sort_gamma_cumul);

GPH = get(handles.GPH,'String');
GPH = str2double(GPH);

gamma_cumul_find = find(sort_gamma_cumul < 1, 1, 'last');
gammaval_1 = (gammaval(x_cumul - gamma_cumul_find) + gammaval(x_cumul - gamma_cumul_find + 1)) / 2;
gammaval_1 = num2str(gammaval_1, '%.2f');

gamma_cumul_find = find(sort_gamma_cumul < 5, 1, 'last');
gammaval_5 = (gammaval(x_cumul - gamma_cumul_find) + gammaval(x_cumul - gamma_cumul_find + 1)) / 2;
gammaval_5 = num2str(gammaval_5, '%.2f');

gamma_cumul_find = find(sort_gamma_cumul < GPH, 1, 'last');
gammaval_GPH = (gammaval(x_cumul - gamma_cumul_find) + gammaval(x_cumul - gamma_cumul_find + 1)) / 2;
gammaval_GPH = num2str(gammaval_GPH, '%.2f');

% Calculate % points with gamma > 1
a = sum(gamma_thresh(:) > 1);
a = double(a);
b = nansum(diff_threshold(:));

pass_rate = 100 - (a / b * 100);
if pass_rate >= 95;
set(handles.gamma_pass,'BackgroundColor',[0 1 0]);

else
    set(handles.gamma_pass,'BackgroundColor',[1 1 1]);
end

pass_rate=num2str(pass_rate,'%.1f');

%Calculate mean and max gamma
mean_g=nanmean(gamma_thresh(:));
mean_g=num2str(mean_g,'%.2f');

median_g=nanmedian(gamma_thresh(:));
median_g=num2str(median_g,'%.2f');

max_g=max(max(gamma_thresh(:)));
max_g=num2str(max_g,'%.2f');

%Display pass rate, mean gamma, and max gamma in GUI
set(handles.gamma_pass,'String',pass_rate);
set(handles.gamma_mean,'String',mean_g);
set(handles.gamma_median,'String',median_g);
set(handles.gamma_max,'String',max_g);
set(handles.gamma_1,'String',gammaval_1);
set(handles.gamma_5,'String',gammaval_5);
set(handles.gamma_GPH,'String',gammaval_GPH);

%--------------------------------------------------
% calculate dose difference using lower and upper cut values
% to be moved from here eventually

dosediffplot_perc=((plane2_threshold-plane_threshold))./plane_threshold)*100;
dosediffplot_abs=(plane2_threshold-plane_threshold).

mean_diff_perc=num2str(nanmean(dosediffplot_perc(:)),'%.2f');
median_diff_perc=num2str(nanmedian(dosediffplot_perc(:)),'%.2f');
std_diff_perc=num2str(nanstd(dosediffplot_perc(:)),'%.2f');

handles.gamma_cumul=gamma_cumul;
handles.gammavalt=gammavalt;
handles.dosediffplot_perc=dosediffplot_perc;
handles.dosediffplot_abs=dosediffplot_abs;

set(handles.mean_diff,'String',mean_diff_perc);
set(handles.median_diff,'String',median_diff_perc);
set(handles.sd_diff,'String',std_diff_perc);

%--------------------------------------------------
% display profile after calculation ends

val4=handles.val4;

%set handle to the lower right axis in the GUI to display profile
axes(handles.gamma_axes);
box on
% display according to the current selection in the popmenu in the analysis area

if val4==1 %Gamma Map
    imagesc(gamma);
    caxis([0 1.6])
    freezeColors
    axis equal
    axis tight
    hold;
    contour(diff_threshold_plot,1,'y');
    hold;
elseif val4==2 %Std Histogram
    hist(gamma_thresh(:))
elseif val4==3 %Cumul histogram
    plot(gammavalt, gamma_cumul)
    axis([0 1.5 0 100])
elseif val4==4 %Difference map(%)
    contourf(dosediffplot_perc)
    colormap(jet)
    freezeColors
    axis equal
    axis tight
elseif val4==5 %Difference map (Gy)
    contourf(dosediffplot_abs)
    colormap(jet)
    freezeColors
    axis equal
    axis tight
elseif val4==6 %DTA map
end
close(h)

% display time taken to perform calculation
message=strcat('Calculation time taken: ',time_taken,' s');
set(handles.time, 'String', message);
handles.gamma_plane=1;
handles.gamma_vol=0;
clc
guida(hObject, handles);
ext2=handles.ext2;
if strcmp(ext2,'.dcm')==1,
    data2_orientation_val=get(hObject,'Value')
    %handles.data2_orientation_val=data2_orientation_val;
    
cube2=handles.cube2;
    
if data2_orientation_val ==1
    cube_plot2=cube2;
elseif data2_orientation_val == 2
    cube_plot2=permute(cube2,[3 2 1]);
elseif data2_orientation_val == 3
    cube_plot2=permute(cube2,[3 1 2]);
end
    max_dose=max(cube2(:));
    handles.cube_plot2=cube_plot2;
    
sizeImg = size(cube_plot2);
minSlice = 1;
maxSlice = sizeImg(3);
startSlice = round((maxSlice-minSlice)/2);
    handles.slice = startSlice;
    
sliderStep = [1, 1] / (maxSlice - minSlice);
    set(handles.slider3,'value',startSlice); %
    set(handles.slider3,'max',maxSlice); %
    set(handles.slider3,'min',minSlice);
    set(handles.slider3, 'SliderStep', sliderStep);
if data2_orientation_val==1
    axes(handles.axes2)
    imagesc(cube_plot2(:,:,startSlice));
colormap(jet)
caxis([0 max_dose])
freezeColors
axis equal
axis tight
else
    axes(handles.axes2)
    imagesc(cube_plot2(:,:,startSlice));
colormap(jet)
caxis([0 max_dose])
freezeColors
axis equal
axis tight
end
    plane2=cube_plot2(:,:,startSlice);
handles.plane2=plane2;
    [sizex,sizey]=size(plane2);
    sizeplane = size(plane2);
minplane = 1;
maxplane = sizeplane(1);
startplane = round((maxplane-minplane)/2);
planeStep = [1, 1] / (maxplane - minplane);

set(handles.profile_scroll,'value',startplane); %
set(handles.profile_scroll,'max',maxplane); %
set(handles.profile_scroll,'min',minplane);
set(handles.profile_scroll, 'SliderStep', planeStep);

axes(handles.profile_axes)
plot(plane2(round(sizex/2),:))

elseif strcmp(ext2,'.xls')==1
    ...
end

guidata(hObject, handles);

info1=handles.info1;
cube1=handles.cube;
cube1_orig=cube1;
cube1=interp3(cube1, interpfactor, interp_text);
if data1_orientation_val ==1 % Axial
    spacing_x_1=info1.PixelSpacing(1)/2^interpfactor;
    spacing_y_1=info1.PixelSpacing(2)/2^interpfactor;
    slice_thickness=info1.GridFrameOffsetVector(2)-
                    info1.GridFrameOffsetVector(1)/2^interpfactor;
elseif data1_orientation_val == 2 % Coronal
    cube1=permute(cube1,[3 2 1]);
    spacing_y_1=info1.PixelSpacing(2)/2^interpfactor;
    slice_thickness=info1.PixelSpacing(1)/2^interpfactor;
    spacing_x_1=(info1.GridFrameOffsetVector(2)-
                    info1.GridFrameOffsetVector(1))/2^interpfactor;
elseif data1_orientation_val == 3
    cube1=permute(cube1,[3 1 2]); % Sagittal
    slice_thickness=info1.PixelSpacing(2)/2^interpfactor;
    spacing_y_1=info1.PixelSpacing(1)/2^interpfactor;
    spacing_x_1=(info1.GridFrameOffsetVector(2)-
                    info1.GridFrameOffsetVector(1))/2^interpfactor;
end

[sizex, sizey, sizez]=size(cube1);
info2=handles.info2;
cube2=handles.cube2;
cube2=interp3(cube2, interpfactor2, interp_text);
if data2_orientation_val ==1 % Axial
spacing_x_2=info2.PixelSpacing(1)/2^interpfactor2;
spacing_y_2=info2.PixelSpacing(2)/2^interpfactor2;

elseif data2_orientation_val == 2 % Coronal
cube2=permute(cube2,[3 2 1]);

spacing_y_2=info2.PixelSpacing(2)/2^interpfactor2;
spacing_x_2=(info2.GridFrameOffsetVector(2)-
info2.GridFrameOffsetVector(1))/2^interpfactor2;

elseif data2_orientation_val == 3

cube2=permute(cube2,[3 1 2]); % Sagittal

spacing_y_2=info2.PixelSpacing(1)/2^interpfactor2;
spacing_x_2=(info2.GridFrameOffsetVector(2)-
info2.GridFrameOffsetVector(1))/2^interpfactor2;

end

[sizex_2, sizey_2, sizez_2]=size(cube2);

search_dist_z = round(distanceCriterion * 4 / (slice_thickness * 
2^interpfactor));
%search_dist=round(distanceCriterion*2*2^interp_1);
search_dist_x=6;
search_dist_y=6;

offset_x=(sizex_2*(abs(spacing_x_2/spacing_x_1))-
((spacing_x_2/spacing_x_1)-1)-
sizex;
offset_y=(sizey_2*(abs(spacing_y_2/spacing_y_1))-
((spacing_y_2/spacing_y_1)-1)-
sizey;

if offset_x > 0

cube2=cube2(1+offset_x/2^interpfactor:sizex_2,
1+offset_y/2^interpfactor:sizey_2, 1:sizez_2);

[sizex_2, sizey_2, sizez_2]=size(cube2);
elseif offset_x < 0 && offset_y < 0

cube2_expand=zeros(sizex, sizey, sizez);

cube2_expand(1:sizex_2, 1:sizey_2, :)=cube2;

cube2=cube2_expand;

[sizex_2, sizey_2, sizez_2]=size(cube2);
end

eval_cube1=zeros(sizex + search_dist_x*2, sizey + search_dist_y*2, sizex + search_dist_z*2);
eval_cube1(:,:,::)=10^8;
eval_cube1(search_dist_x + 1 : search_dist_x + sizex,...
search_dist_y + 1: search_dist_y + sizey, ...
search_dist_z + 1 : search_dist_z + sizez) = cube1;
Appendix C: Gamma index software matlab code extract

gValue=zeros(search_dist_x * 2 + 1, search_dist_y * 2 + 1, search_dist_z * 2 + 1);
gValue(:,:,1)=100;
distdiff=zeros(search_dist_x * 2 + 1, search_dist_y * 2 + 1, search_dist_z * 2 + 1);
gamma=zeros(sizex_2, sizey_2, sizez_2);

for k=1:search_dist_z * 2 + 1
    for i=1:search_dist_x * 2 + 1
        for j=1:search_dist_y * 2 + 1
            distdiff(i,j,k)=sqrt(sqrt(((i-search_dist_x-1)*spacing_x_1)^2+((j-search_dist_y-1)*spacing_y_1)^2)^2+(k-search_dist_z-1)*slice_thickness)^2);
        end
    end
end

%######################## 3D vs 3D volume gamma analysis: Wendling proposal

for slice_num=1:sizez_2
    waitbar(slice_num/sizez_2,h);
    if getappdata(h,"canceling")
        break
    end
    for x=1:sizex_2
        for y=1:sizey_2
            testdose=eval_cube1(1+(x-1)*spacing_x_2/spacing_x_1:(x-1)*spacing_x_2/spacing_x_1+search_dist_x*2+1,...
                                1+(y-1)*spacing_y_2/spacing_y_1:(y-1)*spacing_y_2/spacing_y_1+search_dist_y*2+1,...
                                slice_num*2^interpfactor-search_dist_z+search_dist_z-(2^interpfactor-1): slice_num*2^interpfactor+search_dist_z+search_dist_z-(2^interpfactor-1));
            if local_global==1,
                test_dosediff=(cube2(x,y,slice_num)-testdose)/normDose;
            else
                test_dosediff=(cube2(x,y,slice_num)-testdose)/testdose;
            end
        end
    end
gValue = sqrt((test_dosediff./doseCriterion).^2 + (distdiff./distanceCriterion).^2);

gamma(x, y, slice_num) = min(gValue(:));

end
end
end