QUEENSLAND UNIVERSITY OF TECHNOLOGY

MASTERS THESIS

Modelling Electron Cut-Out Factors

Author:
Mr. Simon Biggs

Supervisors:
Mr. John Kenny
Dr. Konstantin Momot
Dr. Andrew Fielding

Submitted to the Science and Engineering Faculty, Queensland University of Technology, in partial fulfilment of the requirements of the degree of Master of Applied Science (Medical Physics).

11th June 2014
© 2014 by Simon Biggs.

Modelling Electron Cut-Out Factors by Simon Biggs is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-sa/4.0/.
Statement of Authorship

The work contained in this thesis has not been previously submitted for a degree or diploma at any other tertiary educational institution. To the best of my knowledge and belief, this thesis contains no material previously published or written by another person except where due reference is made.

Signature:

Simon Biggs

11th June 2014
Abstract

Within is a method for the use of clinically collected measurements in the modelling and prediction of output factors for electron beams used in radiotherapy. In particular the modelling of the electron insert cut-out factor. The benefit of this method is that it reduces the frequency with which new factors must be measured, while providing quality assurance for newly measured factors. It is also less computationally expensive than other methods such as Monte Carlo; found in the literature. Clinical data collected from Radiation Oncology Queensland (ROQ) over the past six years was used to develop this method. The model required the prediction of unknown output factors, and a measure of the confidence of that prediction. This was achieved using an adapted smoothing spline model to fit against the same parameters used by Turian et al [1] (width and aspect ratio). This is in agreement with the collimator and shielding effects discussed by Faddagen et al [2] and the trend between perimeter and area as discussed by Nair et al [3]. The measure of confidence was evaluated using the flexibility of the fit and comparing how this varied with insert shape and size. The uncertainty was given by the final confidence prediction, which was taken to be the maximum value of the various methods employed (between 0.3% and 2%). A validity analysis of the fit was performed using computational resampling methods. Within the ROQ data sets provided, there were five applicator and energy combinations which contained sufficient data for fitting and validation. Predictions for these were all found to be valid using this method.
Acknowledgements

*In all your ways acknowledge him.* — PROVERBS iii. 6.

I desire to acknowledge he who straightens my path. He who is my helper and comforter.

I would like to acknowledge my beautiful wife Esther. She patiently endured the countless days and nights of me being cooped up in my study and unable to aid in parenting our beautiful little boy.

Next I would like to acknowledge John Kenny, my supervisor and chief medical physicist at ROQ. He took time out of his exceptionally busy schedule, working all over the country, to not only supervise this project but also to mentor me as a work experience medical physicist. I greatly appreciate his patience and willingness to humour my questioning, wade through my emails, and provide me with the preliminary skills required to go on and become a registrar medical physicist.

I would like to thank all of the staff at ROQ who have assisted me in this project either by going out of their way to provide resources, or to answer some my many questions. Worthy of particular mention is Rebecca Murray, who meticulously recorded the majority of the clinical electron data measurements, collated them into electronic format and laid the primary foundations for this project. Furthermore, I would like to thank her for the fact that due to her enthusiastic demonstrations at ROQ of the Medical Physics profession to my medical student friend I confirmed my decision to become a Medical Physicist. Without Rebecca this thesis would not exist.

I wish to acknowledge Konstantin Momot who firstly spent the time sitting with me as I
explained my computational methodology, providing guidance and insight. Secondly he provided continued guidance on the writing of this thesis even while he had an excessive university semester work load.

And finally I wish to acknowledge my friends Matthew Crank, Christopher Bennett, and David Geelan who were able to provide timely feedback on my often unfortunate writing style.
Contents

Statement of Authorship i
Abstract ii
Acknowledgements iii
List of Figures vii
List of Tables xii

1 Background 1
  1.1 Radiation therapy ........................................ 1
  1.1.1 Electron beam radiation treatments ..................... 1
  1.2 Monitor Unit dose calculations for electron beams at ROQ 3
  1.2.1 The cut-out, applicator, and SSD factors .................. 6
  1.3 Electron beam treatment planning .......................... 6
  1.3.1 The use of Varian Eclipse at ROQ ...................... 7
  1.3.2 Pencil beam XiO by Elekta ............................. 8
  1.3.3 The current ROQ output factor interpolation model ......... 8
  1.4 The dependence of dose on cut-out shape .................... 9
  1.4.1 Side scatter equilibrium ................................ 9
  1.4.2 Insert shielding effects ................................ 10
  1.5 Complications with respect to patient geometry ............. 11
  1.6 Electron output factor modelling in the literature 11
  1.6.1 Equivalent field size as a parameter for prediction .... 12
  1.6.2 Monte Carlo ............................................. 13
  1.6.3 Pencil beam ............................................. 13
  1.6.4 Sector integration ....................................... 14
  1.6.5 Justification for the parameters used for prediction .... 15
  1.7 Project Objective .......................................... 15

2 Materials & Methods 16
  2.1 Cubic smoothing spline adapted for three dimensions ........ 16
  2.2 Determining the uncertainty of predictions ................... 16
    2.2.1 Initial fit error ....................................... 18
    2.2.2 Fit give ............................................. 19
    2.2.3 Angle gap ............................................ 22
    2.2.4 Fit slope ............................................ 22
    2.2.5 Calculating the lower bound uncertainty ................. 26
    2.2.6 Flexibility of fit ..................................... 26
    2.2.7 Final uncertainty prediction ........................... 27
## CONTENTS

<table>
<thead>
<tr>
<th>2.3 Computational validation of fit and uncertainty model</th>
<th>27</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.4 Identifying potential outliers</td>
<td>29</td>
</tr>
</tbody>
</table>

### 3 Results

<table>
<thead>
<tr>
<th>3.1 Overview of fits able to be validated</th>
<th>31</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2 Overview of fits unable to be validated</td>
<td>31</td>
</tr>
<tr>
<td>3.3 Potential data outliers</td>
<td>31</td>
</tr>
<tr>
<td>3.4 Model validation</td>
<td>37</td>
</tr>
</tbody>
</table>

### 4 Discussion

<table>
<thead>
<tr>
<th>4.1 Predicting the cut-out factor with known uncertainty</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1.1 Reducing valid region</td>
<td>40</td>
</tr>
<tr>
<td>4.1.2 Adaptation for data sets with small numbers of points</td>
<td>40</td>
</tr>
<tr>
<td>4.1.3 The use of slope</td>
<td>41</td>
</tr>
<tr>
<td>4.2 Lowering the uncertainty</td>
<td>41</td>
</tr>
<tr>
<td>4.2.1 Outliers</td>
<td>41</td>
</tr>
<tr>
<td>4.2.2 Uncertainty bounds</td>
<td>41</td>
</tr>
<tr>
<td>4.2.3 Model validations</td>
<td>41</td>
</tr>
<tr>
<td>4.3 Effective clinical use</td>
<td>42</td>
</tr>
</tbody>
</table>

### 5 Recommendations for Further Study and Conclusions

<table>
<thead>
<tr>
<th>5.1 Effective clinical implementation</th>
<th>43</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2 Further data collection</td>
<td>43</td>
</tr>
<tr>
<td>5.3 Shape detail</td>
<td>44</td>
</tr>
<tr>
<td>5.4 Possibility for an adapted pencil beam model</td>
<td>44</td>
</tr>
<tr>
<td>5.5 Conclusion</td>
<td>45</td>
</tr>
</tbody>
</table>

### A Proof of parameter relationship

<table>
<thead>
<tr>
<th>A.1 Ellipse</th>
<th>46</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.2 Rectangle</td>
<td>47</td>
</tr>
</tbody>
</table>

### B Primary code used

<table>
<thead>
<tr>
<th>B.1 Program and package requirements</th>
<th>48</th>
</tr>
</thead>
<tbody>
<tr>
<td>B.2 Variable explanations</td>
<td>48</td>
</tr>
<tr>
<td>B.3 The fit</td>
<td>49</td>
</tr>
<tr>
<td>B.3.1 Point prediction</td>
<td>49</td>
</tr>
<tr>
<td>B.3.2 Smoothing spline</td>
<td>50</td>
</tr>
<tr>
<td>B.4 The region parameters</td>
<td>51</td>
</tr>
<tr>
<td>B.4.1 Angle gap</td>
<td>51</td>
</tr>
<tr>
<td>B.4.2 Fit give</td>
<td>52</td>
</tr>
<tr>
<td>B.4.3 Fit slope</td>
<td>53</td>
</tr>
</tbody>
</table>

### Bibliography

| Bibliography                                           | 54 |
List of Figures

1.1 The characteristic PDD for an electron beam. ........................................... 2
1.2 An example of a variety of percentage depth dose curves [4]. As can be seen for electron energies with the most useful range (~6-13 MeV) there is a sharp drop off in dose with respect to depth. ........................................... 3
1.3 The collimation system used to conform the electron beam to the target volume. The output of the machine is defined by the measured charge in the ion chamber. The conversion from machine output to dose is initially defined under reference scatter conditions. However this machine output to dose conversion varies depending on the position of the jaws, the size of the applicator, and the size and shape of the final field defining aperture. A total output factor needs to be included in the machine output to dose conversion whenever the set up is not under reference conditions. ........................................... 4
1.4 An applicator attached to the linac head. The purpose of the applicator is to allow the conformation of the electron beam close to the skin surface. Its design is to minimise the leakage of scatter. Picture by Vishal Majithia, used with permission. ........................................... 5
1.5 An insert within an applicator. Its shape is dependent upon the shape of the cancer being bombarded. The material used to create the shape within the insert is a low melting point lead alloy known as Cerrobend\textsuperscript{TM}. Picture by Vishal Majithia, used with permission. ........................................... 5
1.6 A flow chart of the electron beam treatment planning procedure at ROQ. .... 7
1.7 An example of a polynomial fit with insert cut-out factors against the equivalent field size calculated using length and width as in Equation 1.3. ......................... 9
1.8 The primary electron beams that reach the detector at point P tend to hit the water above the detector within a circle of radius $\approx 2$ cm (depending on energy). The bremsstrahlung, photon scatter, and electron leakage as a result of the shielding effects add on to this primary beam causing a dose that is $1 - 2\%$ higher than that which would be expected from the primary beam alone. The bremsstrahlung and photon scatter depend strongly on the amount of shielding used and the electron scatter and leakage through the internal section of the cut-out depend highly on the cut-out’s perimeter. ........................................... 10
1.9 Demonstration of the loss of side-scatter equilibrium. Redrawn from Khan et al. [5]. In this figure a Gaussian probability distribution is drawn over the top of a cross section of the insert, this probability distribution represents the probability that an electron exiting the insert will reach the point of depth maximum along the central vertical axis ($R_{100} / d_{\text{max}}$). ........................................... 11
1.10 An extreme case demonstrating the effect of sharp surface irregularities. [4] ... 12
2.1 A single cubic smoothing spline is chosen from the fit. The chosen spline is drawn in red. It is taken for the arbitrary aspect ratio of 0.66. See Figures 2.2 and 2.3 for a demonstration of the weighting with regards to this particular spline.

2.2 A single spline fit extracted from the spline mesh as shown in Figure 2.1. The weighting given each point is determined from the aspect ratio difference between the data point and that of the spline fit section. Distance weighting is calculated using a normal distribution which can be seen in Figure 2.3. This weighting is represented by the colour of the points shown.

2.3 A plot of an example Guassian distribution used to weight the points for the slice of the three dimensional spline fit shown in Figures 2.1 and 2.2. This slice is at an aspect ratio of 0.66.

2.4 Two points which would presumably have differing confidence in prediction due to the number of data points in the vicinity of the prediction region. A representation of the $\alpha_{+}$ for these points is shown for the higher and lower confidence point in Figures 2.5 and 2.6 respectively.

2.5 A representation of the process of calculating the $\alpha_{+}$ at the point given by a width of 6.3 cm and aspect ratio of 0.8. As can be seen, even though an extreme point is included in the fitting process the fit only moves up to meet the point by 0.10. This represents a $\alpha_{+}$ of 0.10 and indicates that there are many other data points apart from the extreme point in close vicinity to that location.

2.6 A representation of the process of calculating the $\alpha_{+}$ at the point given by a width of 7.3 cm and aspect ratio of 0.85. As can be seen, the inclusion of an extreme point at this location causes the fit to adjust significantly. The adjustment at that location is 0.27 which represents the $\alpha_{+}$. This indicates that there are not as many data points in the vicinity of this prediction location when compared to Figure 2.5.

2.7 $\alpha_{+}$ on its own is not a good measure for uncertainty. As it can be seen there are regions of outliers that follow predictable patterns. This indicates that there should be a method that allows the removal of these points. Further details of this outlier group is given in Figure 2.8.

2.8 Upon investigation of the highlighted outlier group it can be seen that they are in a localised region. It turns out these outliers have resulted due to extrapolation close to the edge points. A method to identify and remove these points is given in Section 2.2.3.

2.9 An example point given by a width of 4.5 cm and aspect ratio of 0.75. This is demonstrating the method by which the greatest $\angle_{gap}$ is found. As it can be seen in this example the greatest angle between two radially consecutive points is 97°. If the cut off was given as 150° this point would be acceptable by the $\angle_{gap}$ interpolation test.

2.10 An example point given by a width of 7.3 cm and aspect ratio of 0.7. This is demonstrating the method by which the greatest $\angle_{gap}$ is found. As it can be seen in this example the greatest angle between two radially consecutive points is 165°. If the cut off was given as 150° this point would not be able to be used.

2.11 A demonstration of the effect of changing the $\angle_{gap}$ cut off. 150° is what is used for the fitting in this project. The method for determining this $\angle_{gap}$ cut off can be seen in Figure 2.12.
2.12 In this figure it can be seen that once the angle gap is allowed to pass 150° the uncertainty in predicted data points begins to increase. A similar trend is visible in varying degrees for each applicator and energy leading to the angle gap cut off of 150° being used. Any point with an angle gap greater than this is assumed to be untrustworthy extrapolation.

2.13 It can be seen that once the angle gap cut off is used give becomes a reasonable indicator of increasing uncertainty. As give increases the spread of the predicted data also spreads. As a result this is one viable solution for classifying uncertainty regions over the fit.

2.14 An example of the slope calculation. The vectors on the right hand figure represent the slope of the fit. The slope at a particular point is defined as the magnitude of these vectors. Notice that to bring the axes onto a scale that is similar to one another the ratio axis was multiplied by the applicator size.

2.15 As can be seen, the slope is a slight predictor of the spread. The progressive standard deviation of the points is shown as the dotted black line.

2.16 The method of calculating the lower bound uncertainty is demonstrated. The yellow boxes represent the standard deviation of each of the groups of 21 measured data points.

2.17 The scaling of the relative fit flexibility to be above the lower bound measurement uncertainty.

2.18 First part of the flexibility demonstration. Simulated measured data were selected at random from a sine curve with noise. This sine curve can be seen on the left. The resulting cubic spline smoothing fit to these simulated data can be seen on the right.

2.19 Second part of the flexibility demonstration. The first step is to adjust all of the data points by a random normal number, then retake the fit. An example of ten adjusted (noisy) fits is shown on the left. Once this procedure is completed many times, the standard deviation of these adjusted fits can be calculated. The first and second standard deviation are represented on the right by the dashed lines.

2.20 Last part of the flexibility demonstration. In this case, we know the true value of the function being predicted (sine). As such, the difference between the sine curve and the original fit is plotted here on the left against the standard deviation predicted from the flexibility. A histogram of these data can be seen on the right. This is the same representation as the validation that is given in the results section. It can be seen that this representation of the uncertainty is a valid representation as no point beyond two standard deviations occurs greater than predicted by the normal distribution (see Section 2.3).

2.21 An example of a potential outlier data point in the 10 cm applicator fit at 6 MeV.

3.1 An overview of the fit for the 6 cm applicator at 6 MeV is shown on the left. On the right the uncertainty contours for this fit can be seen. It can be seen that the overwhelming majority of the valid area has its uncertainty between 0.010 and 0.015.

3.2 An overview of the fit for the 6 cm applicator at 9 MeV is shown on the left. On the right the uncertainty contours for this fit can be seen. It can be seen that the overwhelming majority of the valid area has its uncertainty between 0.008 and 0.012.
3.3 An overview of the fit for the 10 cm applicator at 6 MeV is shown on the left. On the right the uncertainty contours for this fit can be seen. It can be seen that the overwhelming majority of the valid area has its uncertainty between 0.005 and 0.006. 

3.4 An overview of the fit for the 10 cm applicator at 9 MeV is shown on the left. On the right the uncertainty contours for this fit can be seen. It can be seen that the overwhelming majority of the valid area has its uncertainty between 0.003 and 0.004.

3.5 An overview of the fit for the 15 cm applicator at 9 MeV is shown on the left. On the right the uncertainty contours for this fit can be seen. It can be seen that the overwhelming majority of the valid area has its uncertainty between 0.003 and 0.004.

3.6 An overview of the fit for the 10 cm applicator at 12 MeV. Not enough points were spread out within the valid regions of the fit. If however the size of the valid region was manually reduced the subset region would be able to undergo validation.

3.7 An overview of the fit for the 15 cm applicator at 6 MeV. Not enough points were spread out within the valid regions of the fit. However, potentially a reduced region may be able to be validated.

3.8 An overview of the fit for the 15 cm applicator at 12 MeV. Not enough points were spread out within the valid regions of the fit. It is unlikely without taking more measurements within the regions of the fit that this fit could be validated.

3.9 An overview of the fit for the 15 cm applicator at 16 MeV. Not enough points were spread out within the valid regions of the fit. It is unlikely without taking more measurements within the regions of the fit that this fit could be validated.

3.10 The results of validating the 6 cm applicator for 6 MeV. Completed using random subsets 17 points. On the left a plot of the errors in cut-out factor predictions. The dotted lines represent the $\sigma$ and 2$\sigma$ bounds. The solid black line is the actual $\sigma$ of the plotted points. The data points resulting from the potential outliers identified in Section 3.3 are marked in red. On the right a histogram of the distribution of these data points is given. The red bars are the potential outlier points.

3.11 The results of validating the 6 cm applicator for 9 MeV. Completed using random subsets 17 points. On the left a plot of the errors in cut-out factor predictions. The dotted lines represent the $\sigma$ and 2$\sigma$ bounds. The solid black line is the actual $\sigma$ of the plotted points. On the right a histogram of the distribution of these data points is given.

3.12 The results of validating the 10 cm applicator for 6 MeV. Completed using random subsets 23 points. On the left a plot of the errors in cut-out factor predictions. The dotted lines represent the $\sigma$ and 2$\sigma$ bounds. The solid black line is the actual $\sigma$ of the plotted points. The data points resulting from the potential outliers identified in Section 3.3 are marked in red. On the right a histogram of the distribution of these data points is given. The red bars are the potential outlier points.
3.13 The results of validating the 10 cm applicator for 9 MeV. Completed using random subsets 17 points. On the left a plot of the errors in cut-out factor predictions. The dotted lines represent the $\sigma$ and $2\sigma$ bounds. The solid black line is the actual $\sigma$ of the plotted points. The data points resulting from the potential outlier identified in Section 3.3 are marked in red. On the right a histogram of the distribution of these data points is given. The red bar is the potential outlier points. ............ 39

3.14 The results of validating the 15 cm applicator for 9 MeV. Completed using random subsets 14 points. On the left a plot of the errors in cut-out factor predictions. The dotted lines represent the $\sigma$ and $2\sigma$ bounds. The solid black line is the actual $\sigma$ of the plotted points. On the right a histogram of the distribution of these data points is given. .................. 39
List of Tables

3.1 Data sets for which modelling and validation was possible. ............... 31
3.2 Data sets containing too few data points to create a fit .................. 34
3.3 Data sets containing too few points spread out within the fit for validation. . . . 34
3.4 Top seven candidates for being potential outliers. ......................... 37
Chapter 1

Background

Electron beam treatments are often used in the clinical setting for the treatment of tumours close to the surface of the skin, such as skin and other superficial cancers. They are also often used in combination with photon treatments to treat cancers such as breast, head and neck cancers. Although other methods exist of treating these cancers, electron therapy has the advantage of a sharp drop off in dose with depth, allowing for the sparing of deeper tissue.

To conform the dose to the outlined cancer region (target volume) a collimation system is used (see Figure 1.3 and Section 1.1.1). Variation of elements within this collimator system will result in variations of the dose magnitude and profile. During treatment planning a variation of shape in the required field is accounted for with an insert cut-out factor (see Section 1.2.1). By creating a computational model to predict these cut-out factors the frequency with which new factors are measured can be reduced (see Section 1.3).

1.1 Radiation therapy

Radiation therapy is a means of causing cell death in unwanted cells, often cancerous, without having to surgically operate on a patient. Either cell death or ceasing the reproductive ability of a cell is sufficient. The accepted mechanism by which radiation therapy achieves this is by a DNA strand break. Unfortunately not only the target cells are irradiated; so too are healthy cells. Therefore, the challenge in radiation therapy is minimising healthy cell damage, while maximising target cell damage.

One method of creating this ionising radiation is by using a linear accelerator (linac). These are capable of creating both electron beams and photon beams at multiple energies. Both modes are capable of being shaped to conform to the therapeutic target. The particular linac referred to throughout this thesis is the Varian 21EX Clinac.

1.1.1 Electron beam radiation treatments

Electron beams are advantageous for surface tumours due to their quickly tapering off dose distribution with depth. This is due to charged particle coulomb force interactions between the electrons in the beam and body tissue, both orbital electrons and nuclei. These processes are:

- inelastic ionisation or excitation of atomic electrons,
- inelastic collisions with nuclei producing bremsstrahlung radiation,
- or elastic collisions with both nuclei and atomic electrons.

In water, which has a low atomic number, the energy loss of the electrons is predominantly due to ionisation events. In higher-atomic-number materials such as the lead alloy insert or
1.1. RADIATION THERAPY

Figure 1.1: The characteristic PDD for an electron beam.

the jaws of the linac, unwanted bremsstrahlung radiation becomes more prominent [4]. This process of the electrons readily interacting with materials allows for the electron beam to be shaped with relatively thin amounts of shielding. Due to the low mass of the electron and the high number of interactions with the medium, the electron beam has a high level of lateral scatter. This results in the penumbra of an electron beam being broader than that for photons, as well as the dose distribution spreading out in an approximately Gaussian fashion with depth. Nevertheless, in many cases the challenges resulting from the use of electrons are outweighed by their benefits.

Electron depth dose curves

The primary reason for the use of electrons is their characteristic depth dose curve (see Figure 1.1). The percentage depth dose curves (PDDs) for a range of electron energies can be seen in Figure 1.2. For the most useful energies 6-13 MeV it can be seen that there is a sharp drop off in dose with depth. This allows for high conformation with respect to depth for target regions that are close to the surface. At ROQ the available electron energies are 6 MeV, 9 MeV, 12 MeV, and 16 MeV. An important characteristic of the dose imparted by an electron is that there is a well defined maximum depth for the electrons in the beam, this is the practical range ($R_p$). Below $R_p$ the photon contamination within the beam dominates the dose. An important reference depth on this curve is known as the depth of maximum dose ($d_{max}$). This is labelled in Figure 1.1.

Electron field shaping and the insert cut-out within the collimation system

Due to the high geometric penumbral associated with electron beams any shielding as a means of shaping the beam needs to be applied to the patient close to the surface. As a result, the beam collimation used for photons is insufficient. To conform the dose to the target volume a full collimation system is used (see Figure 1.3). This includes the jaws in the head of the linear accelerator (linac), the applicator attached to the head which scrapes off penumbral electrons (see Figure 1.4) and an insert within the final field defining apparatus (see Figure 1.5). This
insert is made from a lead alloy called Cerrobend™. This is a particular commercial name for “Wood’s metal” or “Lipowitz’s alloy”. The make up of Cerrobend™ by weight is 50% bismuth, 26.7% lead, 13.3% tin, and 10% cadmium [6]. It is primarily the effects of shaping the beam with this insert, within the final field defining aperture, that is focussed on in this thesis.

1.2 Monitor Unit dose calculations for electron beams at ROQ

The primary ionisation chambers within the linac head measure fluence in Monitor Units (MUs). The sensitivity of the chamber electrometer circuitry is adjusted so that 100 MU corresponds to a dose of 1 Gy at \( d_{\text{max}} \) in a water phantom under reference conditions [7]. This reference dose is defined as \( D_{d_{\text{max}}}(F_{\text{ref}}, SSD_{\text{ref}}) \). The electron beam reference conditions are given by a 10 x 10 cm standard insert within a 10 x 10 cm applicator with a source to surface distance (SSD) of 100 cm.

The determination of absorbed dose at a depth of \( d_{\text{max}} \) under non-reference conditions requires the use of an electron output factor, defined as \( OF(F, SSD) \). The non-reference absolute dose is given as \( D_{d_{\text{max}}}(F, SSD) \). The definition of this output factor is given below in Equation 1.1 [5]:

\[
\frac{D_{d_{\text{max}}}(F, SSD)}{MU} = OF(F, SSD) \times \frac{D_{d_{\text{max}}}(F_{\text{ref}}, SSD_{\text{ref}})}{MU}
\]  

(1.1)

Since at ROQ \( D_{d_{\text{max}}}(F_{\text{ref}}, SSD_{\text{ref}})/MU = 1/100 \) (given units of dose is in Gy) the electron output factor is then able to be used in conjunction with the MU to calculate the absolute dose at \( d_{\text{max}} \) Equation 1.2:

\[
D_{d_{\text{max}}}(F, SSD) = \frac{MU}{100} \times OF(F, SSD)
\]  

(1.2)

Once this absolute dose at \( d_{\text{max}} \) has been measured/calculated/predicted relative dose distributions, such as a PDDs and dose profiles, are able to be used to predict the doses at depths...
Figure 1.3: The collimation system used to conform the electron beam to the target volume. The output of the machine is defined by the measured charge in the ion chamber. The conversion from machine output to dose is initially defined under reference scatter conditions. However this machine output to dose conversion varies depending on the position of the jaws, the size of the applicator, and the size and shape of the final field defining aperture. A total output factor needs to be included in the machine output to dose conversion whenever the set up is not under reference conditions.
1.2. MONITOR UNIT DOSE CALCULATIONS FOR ELECTRON BEAMS AT ROQ

Figure 1.4: An applicator attached to the linac head. The purpose of the applicator is to allow the conformation of the electron beam close to the skin surface. Its design is to minimise the leakage of scatter. Picture by Vishal Majithia, used with permission.

Figure 1.5: An insert within an applicator. Its shape is dependent upon the shape of the cancer being bombarded. The material used to create the shape within the insert is a low melting point lead alloy known as Cerrobend™. Picture by Vishal Majithia, used with permission.
other than $d_{\text{max}}$ and positions other than the central axis.

This total electron output factor is a function of SSD, and the field (F). The field varies according to a change in insert cut-out or applicator. As a result the output factor is divided into a cut-out factor, applicator factor, and an SSD factor. These will be discussed further in the following section.

1.2.1 The cut-out, applicator, and SSD factors

The output factor is able to be divided into a cut-out factor, an applicator factor, and an SSD factor. Each of these takes into account a particular form of variation from the reference condition.

**Cut-out factor**

The cut-out and applicator factors take into account changes in the field. Specifically, the cut-out factor takes into account the patient-to-patient variation of the insert cut-out that sits inside the final field shaping aperture. This cut-out defines the primary shape of the beam. The shape and size of this cut-out has a significant effect on the output factor. This is discussed in further detail in Section 1.4. It is the cut-out factor that varies from patient to patient and needs to be predicted/measured/calculated for each treatment.

**Applicator factor**

The applicator factor takes into account not just the applicator effects, but also the effects of the jaws within the head of the linac. For each specific applicator, the jaws have a different pre-defined position. By changing the applicator not only are the scraping effects of the applicator adjusted, but so too are the field effects that result from changing the jaws.

**SSD factor**

The SSD factor is defined using the inverse square law (ISL). At ROQ patients are treated at either 100 cm SSD, 110 cm SSD, or 120 cm SSD. The change in fluence effects that result due to the ISL are taken into account using this factor.

1.3 Electron beam treatment planning

The process of planning an electron treatment is multi-step. An overview of each of these steps is given in the numbered points just below. A flow chart of the procedure at ROQ can be seen in Figure 1.6. In Section 1.3.1 and following, the points of relevance to this thesis will be discussed in more detail.

1. Doctor contours the planning treatment volume (PTV), then prescribes a depth and dose required.

2. The energy of the photon beam is decided so that the tumour is within the 90% dose treatment region of the PDD (see Figure 1.1).

3. The treatment planning system (TPS) creates a cut-out shape that will provide sufficient margin around the PTV in order to account for the electron beam penumbra.

4. At this point a relative dose distribution is calculated in a method that depends upon the specific TPS in use.
5. Using the relative dose distribution, the dose required at the calculation point (e.g. $d_{\text{max}}$ along the central axis) is determined in order to meet the required prescription to the PTV.

6. Finally, a method for calculating/measuring/predicting the output factor (MU required to meet this dose at the calculation point) is employed. This is centre specific, and can be done manually or by the TPS.

### 1.3.1 The use of Varian Eclipse at ROQ

At ROQ the TPS used is Varian Eclipse. To calculate the relative dose distribution, a modified Monte Carlo algorithm is used called eMC (electron Monte Carlo). This same algorithm is able to be employed in order to calculate electron output factors for MU calculation. Historically these commercial systems have not produced output factors with acceptable accuracy. It is for this reason that at ROQ output factors are measured or interpolated from previous measurements, independent of the TPS. The choice of whether or not a factor is to be measured or interpolated is decided on a case-by-case basis. If ever there is any uncertainty, the output factor is measured in RW3 “solid water”.

A flow chart of the output factor procedure at ROQ is given in Figure 1.6. The method of interpolation used is described in further detail in Section 1.3.3. It is this interpolation...
method that this thesis aims to improve, with the purpose of increasing accuracy in output factor predictions and reducing the frequency of RW3 measurements.

1.3.2 Pencil beam XiO by Elekta

Some of the methods for determining output factors outlined in the literature review (see Section 1.6) have been implemented in commercial treatment planning systems. One example is XiO by Elekta, which makes use of the pencil beam model (see Section 1.6.3) for the calculation of the electron output factor. However, as with the eMC algorithm, this too has not historically produced electron output factors within an acceptable accuracy.

1.3.3 The current ROQ output factor interpolation model

ROQ physicists have measured output factors for a large number of patient specific and standard shaped “library” cut-outs. This has occurred for the past six years. The first year’s worth of these measurements have been implemented into the predictive model currently in use. This predictive model is founded upon the idea of an equivalent field size, where complicated field shapes are represented by different yet approximately equivalent fields that are easier to quantify. Two equivalent field definitions are used within this model, given by Equations 1.3 and 1.4:

\[
\text{Equivalent field size}_1 = \sqrt{\text{width} \times \text{length}}
\]

(1.3)

\[
\text{Equivalent field size}_2 = \pi (r_{\text{avg}})^2
\]

(1.4)

The first definition is used to index the measured cut-out factors. The second definition is used to take the information provided by the treatment planning system and convert to a given cut-out factor. This is then able to be combined with the required applicator and SSD factors to calculate the final MU required.

The creation of the current interpolation model

The current model in use was created after one year’s worth of measured cut-out factors had been collected. This library of data was grouped into beam energy and applicator size, then for each energy/applicator combination the cut-out factors were indexed by their equivalent field size as defined by Equation 1.3. To aid in interpolation, curves were fitted to the factors. An example of what this might look like is given in Figure 1.7

The use of the current interpolation model

The current model was implemented to allow the radiation therapists to derive the output factor during treatment planning. This was done by defining a second equivalent field size definition derived from the data available from the TPS. This second definition can be seen in Equation 1.4. The TPS defines its cut-outs by providing 100 shape defining points along the perimeter. To directly convert from this output into an equivalent field size, these shape defining points are used to define an average radius. This average radius is calculated by taking the average of the distance between the shape defining points and the cut-out center. Using Equation 1.4, this average radius is turned into an equivalent field size. Equation 1.4 is based upon the formula for circle area. The obtained equivalent field size is then used to interpolate a cut-out factor from the ROQ model. This cut-out factor is then combined with
1.4. THE DEPENDENCE OF DOSE ON CUT-OUT SHAPE

There are multiple effects at play that contribute to the dose at a given measuring point (see Figure 1.8). When it comes to the shielding used, there are two primary effects discussed by Faddegon et al. [2] which result from high-energy photons being produced in the lead (bremsstrahlung) or electrons scattering from (or through) the inside of the cut-out factor. When it comes to travelling through the medium, the electrons experience coulomb forces primarily due to interactions with nuclei. This results in a ‘random walk’ along the plane perpendicular to the beam. Although primarily the electrons continue in the forward direction, their distribution spreads out laterally with depth in an approximately Gaussian distribution. In practice this lateral spread is no more than 2-3 cm for 99% of the electrons within the primary beam. This actual ‘maximum’ spread can be approximated as $0.66 \times \sqrt{\text{Energy}}$ [4]. For a 6 MeV beam this results in shapes of widths less than 3.2 cm having their output factor dominated by lateral scatter effects, however, for widths greater than 3.2 cm the insert shielding effects would dominate.

1.4.1 Side scatter equilibrium

For small field sizes the dose is highly dependent on the insert shape [5, 8, 9]. This is due to the loss of electron side-scatter/lateral-scatter equilibrium (see Figure 1.9). This probability distribution has an approximate standard deviation of $0.22 \times \sqrt{\text{Energy}}$ [4] and is dependant upon three things: the location of the electron as it passes through the insert, the angular distribution of the electrons at that point, and the probability that the particle will be scattered such that it will arrive at the reference point $d_{\text{max}}$. The implication of this is that in order to have any significant effect on the lateral scatter of the beam, the smallest width of the cut-out would...
1.4. THE DEPENDENCE OF DOSE ON CUT-OUT SHAPE

need to encroach on the third standard deviation bound (the 99% point). This width, at which lateral scatter begins to be lost, is energy dependant and ranges from 3.2 cm for 6 MeV to 5.3 cm at 16 MeV.

1.4.2 Insert shielding effects

The theory behind the effect of cut-out shielding on the output factor was primarily gathered from the output factor predicting Monte Carlo simulations conducted by Faddegon et al. [2]. They found that there were two electron interactions with the shielding that if left unaccounted for could result in underestimates of the output of $\approx 1 - 2\%$. These two effects were labelled the “shielding effect” and the “cut-out collimator effect”. These two effects can explain how cut-outs smaller than the reference insert are able to deposit a greater dose than the reference.

The “cut-out collimator effect” occurs when an electron is scattered out through, or from, the internal surface of the cut-out. This effect approximately increases with increased internal surface area, and therefore increased cut-out perimeter. It is presumed, however, that variations in the shape of this inside surface as a result of irregular cut-outs would produce complicated variations.
1.5 Complications with respect to patient geometry

Throughout this thesis the output factor calculations are predicted using the flat geometry of a water tank or equivalent. The methods described herein allow for prediction uncertainties for this geometry of <1%. It needs to be noted that patient geometry is rarely equivalent to this ideal water tank case. For surface irregularities or inhomogeneities within the patient, the output factors predicted using water tank equivalent methods may have significant systematic errors. An extreme case is given in Figure 1.10. It is for this reason that bolus is used for the purpose of producing a flat perpendicular surface to the electron beam. These patient geometry effects need to be considered when deciding to accept any quoted output factor.

1.6 Electron output factor modelling in the literature

Throughout the literature, modelling of electron factors is done in a variety of ways. Many methods require taking a large number of specific measurements for calibration purposes. Some require computationally intensive approaches while others require highly detailed information about the linac. A number were relatively simple to implement, but only worked for regular shapes. Many work with measured data points, but they often require more detailed information than what has been recorded.
1.6. ELECTRON OUTPUT FACTOR MODELLING IN THE LITERATURE

Figure 1.10: An extreme case demonstrating the effect of sharp surface irregularities. [4]

A method that used arbitrary cut-out factors with only the parameters length and width for predicting both regular and irregular shapes was absent in the literature. The parameters width and aspect ratio were decided upon as a reasonable use of the parameters available for reasons discussed below.

1.6.1 Equivalent field size as a parameter for prediction

Equivalent field methods aim to describe complicated cut-outs in terms of simpler shapes. The method currently in use at ROQ is of this form (see Section 1.3.3). The most well known method of this form for electron output factors is the square root method [5, 10, 11]. It takes output factor measurements from two different squares then combines them in order to find the output factor for a rectangle with length and width defined by each of the squares. This can be seen in Equation 1.5:

\[
OF^{X,Y} = [OF^{X,X} \cdot OF^{Y,Y}]^{1/2}
\]  

(1.5)

Where \(OF^{X,X}\) is the output factor of the square with side \(X\), \(OF^{Y,Y}\) is the output factor of the square with side \(Y\), and \(OF^{X,Y}\) is the desired output factor of the rectangle with width \(X\).
and length \( Y \). It is noted that the square root method should only be applied to rectangles and that even in the case of high aspect ratios this method breaks down \([5]\). Although the work to date at ROQ appears similar to this, it differs in the fact that distance measurements are combined as opposed to the actual output factor measurements being combined in the method discussed here.

Nair et al \([3]\) found that there was a trend between the area/perimeter ratio and the cut-out factor. Their equivalent field size method incorporated this. Unfortunately, area and perimeter are not parameters available for the purpose of the model used in this project. However, the aspect ratio of a shape is available, and is related to the area/perimeter ratio. The trend found between area and perimeter is supported by the theoretical insight provided by Faddegon et al \([2]\) regarding the effects of the shielding (area related) and the insert cut-out collimator effect (perimeter related).

### 1.6.2 Monte Carlo

Many authors discuss the benefits of calculating electron output factors using a Monte Carlo method \([1, 2, 12–16]\). Unique methods are used by each of them to balance computational time and output quality. Two of these methods from which inspiration was drawn for the present project are the Final Aperture Superposition Technique (FAST) by Faddegon et al \([2]\) and the parabola fit of Monte Carlo data by Turian et al \([1]\).

The superposition method used by Faddegon et al \([2]\) involves pre-running the Monte Carlo simulation. Firstly, for an insert that is all shielding and no cut-out, and then secondly, for an insert that is completely empty (or the standard square). Then when a particular output factor is desired, the insert shape is loaded into the program. Any part of the shape that has “Wood’s metal” takes electrons from the pre-run set with the completely full insert. Any part of the insert that is not shielded, selects electrons from the empty insert. These sets of pre-run electrons are then super-positioned on-top of one another to calculate the final dose. This results in a fast and accurate calculation for the output factor, given water tank geometry. Faddagon et al’s \([2]\) results showed a large contribution of bremsstrahlung from shielding and scattered electrons from the cut-out edge, which served as the foundation for the justification of the parameters—width and aspect ratio—used in this model.

A three dimensional parabola fit method was used by Turian et al \([1]\). They created large sets of data via Monte Carlo modelling and then created a 3D parabolic fit. The fit used width and aspect ratio as its parameters. The coefficients of the fit were then recorded and later the fit was used to output the required output factors. Once again, this managed to combine the increased accuracy of Monte Carlo with a method that allowed for fast calculations. The parameters used in this fit—width and aspect ratio—were shown to be valid using their data. These parameters are used in the model of this thesis.

### 1.6.3 Pencil beam

The pencil beam model is primarily a computational method, which will often require the schematics of the linac head to construct. This model is founded on the assumption that the electrons’ lateral spread is in a random walk fashion, resulting in a Gaussian distribution of electrons of increasing spread with depth. There are two primary pencil beam models: the Fermi-Eyges pencil beam, and the lateral build-up ratio (LBR) pencil beam. The difference between the two is the method of predicting/measuring the spread with respect to depth of the Gaussian distribution, \( \sigma_r(z) \). The Fermi-Eyges pencil beam model is implemented within the XiO TPS (see Section 1.3). Historically this model has not produced output factors within the required accuracy. No specific information about the lead alloy shielding is used in pencil
beam model. As a result each of the shielding effects discussed by Faddegon et al. [2] is not accounted for. This by itself would result in prediction errors of $\approx 2\%$. There are also further issues with this modelling method which are beyond the scope of this thesis.

**Fermi-Eyges pencil beam**

The Fermi-Eyges pencil beam model makes use of Fermi’s multiple scattering theory [17] to calculate the spread of the Gaussian with respect to depth, $\sigma_r(z)$. See Equation 1.6:

$$\sigma_r^2(z) = \frac{1}{2} \int \left( \frac{\bar{\theta}^2}{\rho(z')} \right) \rho(z')(z - z')^2 dz'$$

(1.6)

The use of this method for calculating output factors is covered extensively in the literature [18–22].

**Lateral build up ratio pencil beam**

The LBR method of calculating pencil beams differs from the Fermi-Eyges model in that $\sigma_r(z)$ is measured experimentally [8, 23–26]. The LBR is calculated from the measurement of two circular cut-outs and is defined as given in Equation 1.7:

$$LBR(R, z, E) = \frac{D(R, z, E)}{D(R_\infty, z, E)} \frac{\phi_i(R_\infty, E)}{\phi_i(R, E)}$$

(1.7)

Where $\phi_i$ is the incident fluence, $R_\infty$ is the insert that has radius large enough such that lateral scatter equilibrium is reached, $D$ is the dose readings, and $z$ is depth. Using the LBR the lateral spread parameter was calculated as given in Equation 1.8:

$$\sigma_r(z) = \frac{R}{\sqrt{\ln[1/(1 - LBR(R, z))]}$$

(1.8)

This allows for more accurate pencil beam models based upon experimental data.

**1.6.4 Sector integration**

The sector integration method was initially implemented by Clarkson [27] for photons. Authors have shown that this is a reasonable method for the calculation of electron output factors [28–30]. The commissioning of such a system is time consuming. This model is dependent upon having a variety of circle inserts. These inserts begin at a radius of 1 cm then increase by 1 cm up to the maximum circle that fits on each applicator. These inserts need to be measured for every applicator and every energy combination.

The output factor resulting from each circle is then fitted with a smoothing curve such that the output factor for any circular cut-out can be predicted. To calculate a required output factor for any shape, the insert shape is divided into equiangular sectors, with each sector having a known radius. The dose is then calculated sector by sector, assuming that the dose delivered from that sector is equivalent to the same fraction of a circle with the same radius.

Part of the work to date at ROQ on this project is similar to this method, however it is significantly different in two aspects. Firstly, the circles used in this method are measured circles, not approximately equivalent circles. Secondly, this method treats each sector individually, calculating each sector’s dose as a relationship to that individual sector’s radius. ROQ takes an average radius and measures the shape as a whole. This means that the significant advantage
of taking into account fine shape detail that sector integration provides is removed in ROQ’s adapted method.

1.6.5 Justification for the parameters used for prediction

In light of the literature, the parameters chosen to model the cut-out factors against are the widths and aspect ratios of the cut-outs. These parameters are directly available from the provided width and length.

Given that the “shielding effect” is known to decrease with applicator area, and that the “collimator effect” is known to increase with increasing cut-out perimeter (see Section 1.4.2), the ratio given by perimeter/area, used by Nair et al [3], would correlate with these effects. Therefore, taking parameters that are approximately equivalent to perimeter/area would be of value. This is true for the parameters width and aspect ratio. In the case of both the rectangle and the ellipse, the area-perimeter relationship, and the width-aspect ratio relationship is approximated by Equation 1.9:

\[
\frac{\text{Perimeter}}{\text{Area}} \approx \frac{2(\text{Aspect Ratio} + 1)}{\text{Width}}
\]  \hspace{1cm} (1.9)

The proof for this relationship is given in Appendix A.

1.7 Project Objective

The gold standard is taken as the measuring of the electron cut-out factor in a water tank. The aim of this project is to produce a model that predicts this gold standard within a known uncertainty. The model is to have this uncertainty as low as possible with the data available. Ideally this would be less than a standard deviation of 0.5%. Furthermore, the method must feasibly be able to be used quickly, effectively, and without confusion in a clinical situation with the software available at ROQ.
Chapter 2
Materials & Methods

The model required the prediction of unknown output factors within regions where data have previously been measured, and a measure of the confidence of that prediction. This was achieved using an adapted smoothing spline model to fit against width and aspect ratio (see Figure 2.1). The measure of confidence was evaluated using the flexibility of the fit and comparing how this varied with insert shape and size. The uncertainty was given by the final confidence prediction which was taken to be the maximum value of the various methods employed. A validity analysis of the fit was determined using computational resampling methods.

2.1 Cubic smoothing spline adapted for three dimensions

Cubic smoothing splines (smoothed polynomial functions) are designed to be used to smooth data in two dimensions. Adaptation is required to have this method of fitting be applicable in three dimensions. To do this the splines were only calculated in terms of the width and cut-out factors, and each point was weighted according to its aspect ratio. As such each particular aspect ratio created a series of parallel 2D splines forming a 3D surface (see Figures 2.1 and 2.2). The weighting with respect to aspect ratio was calculated using a Gaussian distribution with a standard deviation given by $0.1$. The choice of $0.1$ allows a center point to have every other point encompassed within $3\sigma$. Spline fits have a smoothing parameter ($p$), this was chosen to be $0.7$. The choices of $p$ and $\sigma$ were found by incrementally changing these parameters until the most reasonable fit for each applicator and energy was observed.

2.2 Determining the uncertainty of predictions

For the purpose of predicting the uncertainty in the fit given by a particular cut-out prediction, a multi-step process was employed. This is described as follows:

1. First the fit was remodelled with each individual point removed and predicted, resulting in a given error between the fit and each point. This error was then compared to three separate parameters labelled as give, gap and slope. Each of these parameters quantified qualities of the fit over varying regions (widths and aspect ratios). Give quantified the susceptibility of the region to outliers. Gap quantified how close to the edge of the fit a point was. Slope was the magnitude of the numerically calculated gradient at each point.

2. In order to calculate lower bounds of uncertainty for each parameter, bins of absolute error in prediction were created by stepping one data point at a time, in groups of 21 data points, from the minimum of a given parameter to the maximum. The distribution
2.2. DETERMINING THE UNCERTAINTY OF PREDICTIONS

Figure 2.1: A single cubic smoothing spline is chosen from the fit. The chosen spline is drawn in red. It is taken for the arbitrary aspect ratio of 0.66. See Figures 2.2 and 2.3 for a demonstration of the weighting with regards to this particular spline.

Figure 2.2: A single spline fit extracted from the spline mesh as shown in Figure 2.1. The weighting given each point is determined from the aspect ratio difference between the data point and that of the spline fit section. Distance weighting is calculated using a normal distribution which can be seen in Figure 2.3. This weighting is represented by the colour of the points shown.
2.2. DETERMINING THE UNCERTAINTY OF PREDICTIONS

1. To calculate the flexibility of a fit each of the data points was adjusted by noise (random normal number) and the fit was recalculated. From this adjusted (noisy) fit a point at random was predicted and compared to the original fit. The adjusted fit for each of the point was used to calculate its give, gap and slope. This refitting process was repeated 10,000 times giving a distribution of differences. The distribution was then compared against give, gap and slope. Following the binning process with 1000 data point bins, the standard deviations were calculated producing flexibility distributions in terms of each parameter.

2. These flexibility distributions were made to be on the order of measurement uncertainty by scaling them up to always be above the minimum bounds.

3. When a point was predicted, its uncertainty was predicted by taking the maximum of the predicted uncertainties using give, gap, slope and its flexibility at that point.

Each step in this process is elaborated further in the following sections.

2.2.1 Initial fit error

To have an initial relationship determining how well the fit would have been able to predict each of the available data points, the error between the fit and the measured data points was calculated. This was done in two ways. First, the data point in question was removed from the fit and the fit was remodelled in order to predict that point. Second, if when the data point was removed it fell outside of the valid region (defined by gap in Section 2.2.3) it was simply compared to the fit without remodelling, seeing as that particular point would have not been able to be predicted. Once these fit differences to the measured data points were calculated the method described in Section 2.2.5 was used to develop a relationship between the distribution of the standard deviations of these bins against parameter values formed the lower bound uncertainties.
2.2. DETERMINING THE UNCERTAINTY OF PREDICTIONS

2.2.2 Fit give

It was presumed that the greater the number of points in the vicinity of the prediction region, the greater the confidence that should be possible in the quoted prediction. An example of two points that are presumed to have comparatively varied confidence are shown in Figure 2.4.

The fit give test was devised to determine the influence of the data on the fit at that point. Give was defined as the ratio amount a fit moves to meet an extreme point placed above or below the fit. The more data points being used to directly predict that position, the greater the confidence in that prediction. This give was calculated by including an extreme positive (or negative) point at the position in question and then recalculating the fit. The extreme point was chosen by increasing (or decreasing) the predicted cut-out factor by 1. An example of a relatively low give of 0.10 is shown in Figure 2.5 which can be compared to a higher give of 0.27 shown in Figure 2.6. Once two separate fits have been calculated—one with an extreme point in the positive direction (give_+), the other with an extreme point in the negative direction (give_-)—the difference in prediction values is taken and the largest difference of the two is taken as the overall give at that point.

Using give alone as a measure of the inherent uncertainty in the fit is not sufficient. This is demonstrated in Figure 2.7. Although it can be seen that as the give increases so too does the uncertainty in prediction, there are still patterns of outliers which should be able to be removed in other ways. This becomes more obvious in Figure 2.8, where the entire section of most predominant outliers is localised in a single region of the fit. It turns out that this region occurs as a result of predicting points that are closely outside the boundary of the set of data points. A method to define this boundary and only predict points within it is required. The method used to achieve this was angle gap, as discussed below.
2.2. DETERMINING THE UNCERTAINTY OF PREDICTIONS

Figure 2.5: A representation of the process of calculating the give at the point given by a width of 6.3 cm and aspect ratio of 0.8. As can be seen, even though an extreme point is included in the fitting process the fit only moves up to meet the point by 0.10. This represents a give+ of 0.10 and indicates that there are many other data points apart from the extreme point in close vicinity to that location.

Figure 2.6: A representation of the process of calculating the give at the point given by a width of 7.3 cm and aspect ratio of 0.85. As can be seen, the inclusion of an extreme point at this location causes the fit to adjust significantly. The adjustment at that location is 0.27 which represents the give+. This indicates that there are not as many data points in the vicinity of this prediction location when compared to Figure 2.5.
2.2. DETERMINING THE UNCERTAINTY OF PREDICTIONS

Figure 2.7: Give on its own is not a good measure for uncertainty. As it can be seen there are regions of outliers that follow predictable patterns. This indicates that there should be a method that allows the removal of these points. Further details of this outlier group is given in Figure 2.8.

Figure 2.8: Upon investigation of the highlighted outlier group it can be seen that they are in a localised region. It turns out these outliers have resulted due to extrapolation close to the edge points. A method to identify and remove these points is given in Section 2.2.3.
2.2. DETERMINING THE UNCERTAINTY OF PREDICTIONS

2.2.3 Angle gap

It is important that the fit used only ever interpolates. Due to the complexity of the shielding and collimator effect relationship, extrapolation would most likely result in nonsensical outputs. As such, a method for determining the cut off of interpolation in three dimensional space needed to be devised. This was achieved by measuring the maximum angle formed between the point in question and two angularly consecutive data points. This maximum angle was defined as angle gap. This was done using the parameters aspect ratio and width/applicator size. The reasoning behind scaling down the width by the applicator size is to ensure that the two parameters were on a similar order of magnitude to each other so that the angle measurements were meaningful. A demonstration of how two such points have their angle gap calculated is given in Figures 2.9 and 2.10. In Figure 2.11 the resulting regions that become acceptable when the angle gap cut off is varied is shown for 150° and 200°. To determine an appropriate angle gap cut off each applicator and energy combination had its gap plotted against the difference between fits acquired from bootstrapping. An angle was chosen that resulted in the removal of the most outliers with minimal sacrifice of likely valid regions. One such plot used in this determination can be seen in Figure 2.12. When the angle gap cut off is taken into account the use of give becomes significantly more valid. This can be seen in Figure 2.13.

2.2.4 Fit slope

The final parameter used to predict regions of higher uncertainty was the localised gradient, or the fit slope. Numerical differentiation was used to calculate the slope in both the x and y directions. The magnitudes of these gradient vectors were then used to define the slope at that point. In order to have both width and ratio equally effect the slope calculation the ratio was multiplied by the applicator size. As a result, for the 10 × 10 applicator the final slope was quoted as the rate of change of cut-out factor with respect to 1 cm of width and 0.1 of the ratio. An example set of slope results is given in Figure 2.14.
2.2. DETERMINING THE UNCERTAINTY OF PREDICTIONS

Figure 2.10: An example point given by a width of 7.3 cm and aspect ratio of 0.7. This is demonstrating the method by which the greatest angle gap is found. As it can be seen in this example the greatest angle between two radially consecutive points is 165°. If the cut off was given as 150° this point would not be able to be used.

Figure 2.11: A demonstration of the effect of changing the angle gap cut off. 150° is what is used for the fitting in this project. The method for determining this angle gap cut off can be seen in Figure 2.12.
2.2. DETERMINING THE UNCERTAINTY OF PREDICTIONS

Figure 2.12: In this figure it can be seen that once the angle gap is allowed to pass 150° the uncertainty in predicted data points begins to increase. A similar trend is visible in varying degrees for each applicator and energy leading to the angle gap cut off of 150° being used. Any point with an angle gap greater than this is assumed to be untrustworthy extrapolation.

Figure 2.13: It can be seen that once the angle gap cut off is used give becomes a reasonable indicator of increasing uncertainty. As give increases the spread of the predicted data also spreads. As a result this is one viable solution for classifying uncertainty regions over the fit.
2.2. DETERMINING THE UNCERTAINTY OF PREDICTIONS

Figure 2.14: An example of the slope calculation. The vectors on the right hand figure represent the slope of the fit. The slope at a particular point is defined as the magnitude of these vectors. Notice that to bring the axes onto a scale that is similar to one another the ratio axis was multiplied by the applicator size.

Figure 2.15: As can be seen, the slope is a slight predictor of the spread. The progressive standard deviation of the points is shown as the dotted black line.
2.2. DETERMINING THE UNCERTAINTY OF PREDICTIONS

2.2.5 Calculating the lower bound uncertainty

To develop a relationship between the parameters (give, gap, and slope) and the distribution of differences found in Section 2.2.1, the standard deviation of subsets of the distribution was desired. This was achieved by binning these differences by stepping one data point at a time, in groups of 21 data points, from the minimum of a given parameter to the maximum. The distribution of the standard deviations of these bins against parameter values formed what was defined as the lower bound uncertainties (see Figure 2.16). Once these lower bounds were found for the areas of the fit where measured points existed, an extension of this into the regions where no measured points were available was required. This was done following a similar binning method used here, except that the differences calculated were from noise adjusted fits. This is discussed in the following section.

2.2.6 Flexibility of fit

Once the uncertainty had been measured at the points of the fit for which observed data points were available, this uncertainty needed to be extrapolated to regions where no data points were found. This was achievable by calculating the relative flexibility of the fit itself. By adding noise to all of the data, the resulting fit will deviate more in some areas than in other areas. The relationship between this relative fit flexibility and each of the parameters (give, gap, and slope) can then be calculated using the binning method described in Section 2.2.5. This relative fit flexibility is then able to be scaled such that this relationship is never less than the lower bounds of uncertainty defined directly from the measured points (see Figure 2.17). In this way, with enough noisy data points, a smooth curve for uncertainty was able to be calculated in terms of

Figure 2.16: The method of calculating the lower bound uncertainty is demonstrated. The yellow boxes represent the standard deviation of each of the groups of 21 measured data points.

In Figure 2.15 it can be seen that as expected, an increase in the slope generally results in a greater increase in the spread of the predicted data points. These points were calculated by taking the difference between successive fits in a bootstrap method.
2.3 Computational validation of fit and uncertainty model

To validate each fit and the uncertainty prediction the entire fitting procedure was repeated with a number of measured data points removed. The spread of removed data points was then compared to the predicted spread for the given parameters. This was repeated with multiple sets of data points and a varied number of data points. A valid model was defined as errors of
2.3. COMPUTATIONAL VALIDATION OF FIT AND UNCERTAINTY MODEL

Figure 2.18: First part of the flexibility demonstration. Simulated measured data were selected at random from a sine curve with noise. This sine curve can be seen on the left. The resulting cubic spline smoothing fit to these simulated data can be seen on the right.

Figure 2.19: Second part of the flexibility demonstration. The first step is to adjust all of the data points by a random normal number, then retake the fit. An example of ten adjusted (noisy) fits is shown on the left. Once this procedure is completed many times, the standard deviation of these adjusted fits can be calculated. The first and second standard deviation are represented on the right by the dashed lines.
greater than $2\sigma$ occurring to an extent no greater than that which would be expected by the Gaussian distribution.

2.4 Identifying potential outliers

Since the measured data used were assumed to be the gold standard, if a point significantly deviated from the majority of points it was assumed that cut-out factors for that applicator and energy were able to display that level of uncertainty. Given the methodology and the number of points in each data set, very few gold standard points should have an uncertainty for prediction that is greater than $2\sigma$.

To identify these potential outliers two tests were performed. One test was of the final fit, and the final test was with only the point in question removed. In both of these fits, the greatest variation when normalised by the predicted $\sigma$ was recorded. Then each measured data point was sorted, from greatest normalised variation to least. The higher a point is on the list, the more likely it is an outlier. Nevertheless it is plausible that none of the data points are outliers. An example of one of the potential outliers is given in Figure 2.21.
Figure 2.21: An example of a potential outlier data point in the 10 cm applicator fit at 6 MeV.
Chapter 3

Results

3.1 Overview of fits able to be validated

Five of the applicator/energy combinations had a sufficient number of points for fitting and validation. These were the 6 cm applicator at 6 MeV and 9 MeV, the 10 cm applicator at 6 MeV and 9 MeV, and the 15 cm applicator at 9 MeV. The resulting fits and predicted uncertainties of these data sets can be seen respectively in Figures 3.1, 3.2, 3.3, 3.4 and 3.5. A summary of these data sets is given in Table 3.1.

3.2 Overview of fits unable to be validated

Of the data sets, those which had less than 12 points were unable to be fitted. Those which primarily had points on the out-skirts of the fit were unable to be validated due to the fact that when any of these points were removed that testing position was no longer within the valid region.

The data sets which had too few points to fit are tabulated in Table 3.2. Until more data is available for these data sets, the fit itself is not able to be formed. Validation on top of that would also require more data, spread out within the region of the fit.

A table of the data sets where there were enough points for creating the fit, but not enough for validation is given in Table 3.3. An overview of each of these fits is given in Figures 3.6, 3.7, 3.8, and 3.9.

3.3 Potential data outliers

The top seven candidates for being potential outliers taken from the fits able to be validated are given in Table 3.4.

<table>
<thead>
<tr>
<th>Applicator (cm)</th>
<th>Energy (MeV)</th>
<th>Number of data points</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>6</td>
<td>37</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>24</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>45</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>22</td>
</tr>
<tr>
<td>15</td>
<td>9</td>
<td>17</td>
</tr>
</tbody>
</table>

Table 3.1: Data sets for which modelling and validation was possible.
3.3. POTENTIAL DATA OUTLIERS

Figure 3.1: An overview of the fit for the 6 cm applicator at 6 MeV is shown on the left. On the right the uncertainty contours for this fit can be seen. It can be seen that the overwhelming majority of the valid area has its uncertainty between 0.010 and 0.015.

Figure 3.2: An overview of the fit for the 6 cm applicator at 9 MeV is shown on the left. On the right the uncertainty contours for this fit can be seen. It can be seen that the overwhelming majority of the valid area has its uncertainty between 0.008 and 0.012.
3.3. POTENTIAL DATA OUTLIERS

Figure 3.3: An overview of the fit for the 10 cm applicator at 6 MeV is shown on the left. On the right the uncertainty contours for this fit can be seen. It can be seen that the overwhelming majority of the valid area has its uncertainty between 0.005 and 0.006.

Figure 3.4: An overview of the fit for the 10 cm applicator at 9 MeV is shown on the left. On the right the uncertainty contours for this fit can be seen. It can be seen that the overwhelming majority of the valid area has its uncertainty between 0.003 and 0.004.
3.3. POTENTIAL DATA OUTLIERS

Figure 3.5: An overview of the fit for the 15 cm applicator at 9 MeV is shown on the left. On the right the uncertainty contours for this fit can be seen. It can be seen that the overwhelming majority of the valid area has its uncertainty between 0.003 and 0.004.

<table>
<thead>
<tr>
<th>Applicator (cm)</th>
<th>Energy (MeV)</th>
<th>Number of data points</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>16</td>
<td>7</td>
</tr>
<tr>
<td>20</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>20</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>20</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>20</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>25</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>25</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>25</td>
<td>16</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 3.2: Data sets containing too few data points to create a fit.

<table>
<thead>
<tr>
<th>Applicator (cm)</th>
<th>Energy (MeV)</th>
<th>Number of data points</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td>15</td>
<td>16</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 3.3: Data sets containing too few points spread out within the fit for validation.
3.3. POTENTIAL DATA OUTLIERS

Figure 3.6: An overview of the fit for the 10 cm applicator at 12 MeV. Not enough points were spread out within the valid regions of the fit. If however the size of the valid region was manually reduced the subset region would be able to undergo validation.

Figure 3.7: An overview of the fit for the 15 cm applicator at 6 MeV. Not enough points were spread out within the valid regions of the fit. However, potentially a reduced region may be able to be validated.
3.3. POTENTIAL DATA OUTLIERS

Figure 3.8: An overview of the fit for the 15 cm applicator at 12 MeV. Not enough points were spread out within the valid regions of the fit. It is unlikely without taking more measurements within the regions of the fit that this fit could be validated.

Figure 3.9: An overview of the fit for the 15 cm applicator at 16 MeV. Not enough points were spread out within the valid regions of the fit. It is unlikely without taking more measurements within the regions of the fit that this fit could be validated.
3.4 MODEL VALIDATION

The results from each of the validation computations can be seen in Figures 3.10, 3.11, 3.12, 3.13, and 3.14. Of the five fits required for validation only the 10 cm applicator at 9 MeV came close to being invalid. All five fits passed validation.

It is important to note that the requirement for validation was not that the histogram distribution was normally distributed. Instead, it was important that the calculated uncertainty was consistently greater than or equal to the actual uncertainty. In each of these validation figures the normal distribution curve is for illustrative purposes. The important parameter is that at no point are the bins representing values outside of the threshold (2 standard deviations) higher than that would be expected by the normal distribution.

For the validation found in Figure 3.14, this validation relied on reducing the 17 available points down to only 14 points. As such, any validation modelling was being done with only a 14 point fit. Even so, given this small number of points, no prediction broke the validation threshold.

### Table 3.4: Top seven candidates for being potential outliers.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Width (cm)</th>
<th>Aspect ratio</th>
<th>Identifier</th>
<th>Standard deviation from fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 cm / 6 MeV</td>
<td>3.05</td>
<td>0.73</td>
<td>3</td>
<td>2.0</td>
</tr>
<tr>
<td>6 cm / 6 MeV</td>
<td>3.75</td>
<td>0.80</td>
<td>20</td>
<td>2.0</td>
</tr>
<tr>
<td>6 cm / 6 MeV</td>
<td>3.10</td>
<td>0.86</td>
<td>5</td>
<td>1.9</td>
</tr>
<tr>
<td>10 cm / 6 MeV</td>
<td>5.60</td>
<td>0.62</td>
<td>98</td>
<td>1.8</td>
</tr>
<tr>
<td>10 cm / 9 MeV</td>
<td>4.00</td>
<td>0.53</td>
<td>127</td>
<td>1.8</td>
</tr>
<tr>
<td>10 cm / 6 MeV</td>
<td>3.70</td>
<td>0.58</td>
<td>82</td>
<td>1.7</td>
</tr>
<tr>
<td>10 cm / 6 MeV</td>
<td>5.60</td>
<td>0.63</td>
<td>97</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Figure 3.10: The results of validating the 6 cm applicator for 6 MeV. Completed using random subsets 17 points. On the left a plot of the errors in cut-out factor predictions. The dotted lines represent the $\sigma$ and $2\sigma$ bounds. The solid black line is the actual $\sigma$ of the plotted points. The data points resulting from the potential outliers identified in Section 3.3 are marked in red. On the right a histogram of the distribution of these data points is given. The red bars are the potential outlier points.
3.4. MODEL VALIDATION

Figure 3.11: The results of validating the 6 cm applicator for 9 MeV. Completed using random subsets 17 points. On the left a plot of the errors in cut-out factor predictions. The dotted lines represent the $\sigma$ and $2\sigma$ bounds. The solid black line is the actual $\sigma$ of the plotted points. On the right a histogram of the distribution of these data points is given.

Figure 3.12: The results of validating the 10 cm applicator for 6 MeV. Completed using random subsets 23 points. On the left a plot of the errors in cut-out factor predictions. The dotted lines represent the $\sigma$ and $2\sigma$ bounds. The solid black line is the actual $\sigma$ of the plotted points. The data points resulting from the potential outliers identified in Section 3.3 are marked in red. On the right a histogram of the distribution of these data points is given. The red bars are the potential outlier points.
3.4. MODEL VALIDATION

Figure 3.13: The results of validating the 10 cm applicator for 9 MeV. Completed using random subsets 17 points. On the left a plot of the errors in cut-out factor predictions. The dotted lines represent the $\sigma$ and $2\sigma$ bounds. The solid black line is the actual $\sigma$ of the plotted points. The data points resulting from the potential outlier identified in Section 3.3 are marked in red. On the right a histogram of the distribution of these data points is given. The red bar is the potential outlier points.

Figure 3.14: The results of validating the 15 cm applicator for 9 MeV. Completed using random subsets 14 points. On the left a plot of the errors in cut-out factor predictions. The dotted lines represent the $\sigma$ and $2\sigma$ bounds. The solid black line is the actual $\sigma$ of the plotted points. On the right a histogram of the distribution of these data points is given.
Chapter 4

Discussion

4.1 Predicting the cut-out factor with known uncertainty

The first objective of this project was to predict the clinically measured data within a known uncertainty. This was possible for the data sets given by the following applicator/energy combinations: 6 cm/6 MeV, 6 cm/9 MeV, 10 cm/6 MeV, 10 cm/9 MeV, and 15 cm/9 MeV. An overview of these fits and their uncertainties was given in Figures 3.1, 3.2, 3.3, 3.4, and 3.5. For each of these fits it can be seen that over the inside region there are sufficient points with their positions spread over the region. The overall trend of the fits correlate with the scatter and shielding effects discussed in Sections 1.4.1 and 1.4.2.

4.1.1 Reducing valid region

There were two fits for which further reduction of the valid region would feasibly result in the ability to validate these models. These were the 10 cm/12 MeV and 15 cm/6 MeV data sets for which an overview is given in Figures 3.6 and 3.7. It can be seen that even though there is a reasonable number of points in each of these cases there are extended regions inside the declared valid regions that have no measured points. It is important to note that even though there may be points surrounding a region, during validation if one of these surrounding points is removed and the fit is recalculated it will fall outside the valid region of the new fit. This is due to the fact that the only reason the valid region extends to a point that lies on the surrounding edge is because that point is there in the first place. Therefore, only internal points are useful in the validation process. It is for this reason that these data sets are unable to be validated over the given region. An example of how the region could be reduced can be seen in the case of the 10 cm/12 MeV data set given in Figure 3.6. It is feasible that if the points below a ratio of 0.65 and the points above a ratio of 0.85 were removed during the gap calculation the resulting region would be able to be validated.

There were two other data sets for which fitting was possible but validation was not. These were the 15 cm/12 MeV and 15 cm/16 MeV fits. They can be seen in Figures 3.8 and 3.9. Even though these data sets had enough data points, there are very few data points within the fits. They would very likely require more measurements spread out within the valid region in order to be able to computationally validate any significant section.

4.1.2 Adaptation for data sets with small numbers of points

Without method adjustment, any data set that had fewer than 21 points was unable to have its measurement lower bound uncertainty calculated as discussed in Section 2.2.5. To overcome this, the fact that increasing either applicator size or energy resulted in a reduction of
uncertainty was able to be used. Therefore, for data sets with fewer than 21 data points the measurement uncertainty calculated from either a smaller applicator or lower energy was used as this was known to be an upper bound. The focus of this method was on data sets with greater than 13 data points. Due to insufficient investigation being completed for data sets with fewer points, all data sets with less than 13 data points were considered unable to be fitted for the purpose of this project. It is possible that very small regions for these fits may be valid, but further investigation would be required. Those data sets that were considered to have too few data points for fitting are listed in Table 3.2.

### 4.1.3 The use of slope

Slope was not a reasonable predictor of uncertainty in its current form. Slope would be significantly more effective in simply propagating the human measurement uncertainty into the final result. If the implementation of this model requires the human measurement of the length and width of the shape, then the use of slope to propagate this uncertainty would be a better approach.

### 4.2 Lowering the uncertainty

The second objective of this project was to render the uncertainty of prediction as small as possible with the available data. To offer significant confidence in the quoted uncertainty being an upper bound of the true uncertainty, it is likely that the predicted uncertainty is often an overestimate. Nevertheless, by providing an overestimate some of the time, when cut-out factors are of a similar variation to the quoted potential outliers they will still be within the uncertainty bounds. Therefore, it may be possible that the uncertainty could be made smaller using the data available. In particular, should the potential outliers be remeasured, and found to be outliers, some uncertainty bounds will be able to be significantly reduced.

#### 4.2.1 Outliers

The top seven candidates for being potential outliers were given in Table 3.4. This list only includes potential outliers from the valid data sets. A full list was collated in order of priority, and as a further investigation a set of these data points was remeasured in a water tank.

#### 4.2.2 Uncertainty bounds

For clinical usefulness it is preferred that the predicted cut-out factors be quoted to within an uncertainty of 0.5%. This would allow for that particular factor to have a prediction error of no more than 1%, for 95% of the time. Both of the 9 MeV fits for the 10 cm and 15 cm applicators were consistently within these bounds. Also, the majority of the valid region for the 10 cm applicator at 6 MeV was within this bound. However, both of the 6 cm applicator fits had uncertainties greater than the desired limit of 0.5%. The 6 cm applicator at 9 MeV had primarily between 0.8% and 1% for its uncertainty. The 6 MeV had the majority of its region between 1% and 1.3%. For these final two fits with the 6 cm applicator the decision of whether or not to use the predicted cut-out factors depends upon the clinically allowable uncertainty.

#### 4.2.3 Model validations

Figures representing the data for the validation of each of the 5 models can be seen in Figures 3.10, 3.11, 3.12, 3.13, and 3.14. It can be seen that as a result of each of these validations
the frequency of data points that fell outside $2\sigma$ never went above that which would be expected by the Gaussian distribution for that given uncertainty and deviation. What is of further interest, is the fact that in the majority case the points which were at the extreme of the validation data were those that had been pre-highlighted as potential outliers in Section 3.3. As mentioned before, this does not confirm that they are outliers, it is possible that the cut-out factors do deviate on that order in those regions. If they are outliers, their remeasuring could result in the ability to quote significantly lower uncertainty predictions.

4.3 Effective clinical use

The third and final objective was to have the fit be able to be used quickly, effectively and without confusion in a clinical situation with the programs that are available at ROQ. In its current form this is not the case. However, the fits themselves and the uncertainty predictions can be easily and readily converted into spreadsheet tables that are able to be implemented with automatic interpolation between adjacent cells.
Chapter 5

Recommendations for Further Study and Conclusions

5.1 Effective clinical implementation

A few steps are required in order to allow the fit described in this thesis to be used quickly, effectively and without confusion in a clinical setting. The first step is to output all of the valid fits into spreadsheet format. This spreadsheet would be able to initially receive width and aspect ratio/length as an input. Then, via interpolation between cells, it would output the cut-out factor prediction and the uncertainty in prediction.

Given this spreadsheet further improvement is still possible. For example, at ROQ cut-outs produced through the TPS are defined using 100 equiangular points. It would be of significant value to have a method for defining the width and aspect ratio using these 100 points. A possible avenue for doing this would be by initially defining the width as the smallest diameter through the central axis. This is due to the fact that this smallest diameter is of high significance to the cut-out factor. Then in order to calculate the aspect ratio Equation 5.1 can be used. This is assuming that the shape is either approximately an oval, rectangle or a progression between the two.

\[
\text{Aspect ratio} \approx \frac{\text{Perimeter}}{\text{Area}} \times \frac{\text{Width}}{2} - 1 \tag{5.1}
\]

In this way, the data from the TPS would be able to make consistent output factor predictions. Furthermore, defining aspect ratio in terms of perimeter and area is useful due to the fact that perimeter/area has been shown to be a good parameter for the prediction of output factors [3] as well as in agreement with the collimator and shielding theory given in Section 1.4.2. Thus, the spreadsheet could be updated to automatically calculate output factors when the 100 treatment planning data points are provided to it. This would allow for the development of a quick, effective and uncomplicated procedure for automatically calculating output factors in the clinic.

5.2 Further data collection

By taking further data measurements the cut-out factor models themselves would be able to be significantly improved. This would potentially reduce uncertainty, increase regions of validity, and increase the number of valid applicator/energy combinations.

In order to potentially reduce the uncertainty, the highest priority group of outliers should
be remeasured in a water tank. If the new measurements find that these are actually outliers, then the fits can be remodelled with the adjusted data. This is likely to significantly reduce the determined prediction uncertainty. Furthermore, if the outliers are confirmed, statistically regressing through the outliers with respect to potential causes would aid in the prevention of these in the future.

Taking extra measurements for the four applicator/energy data sets that were unable to be validated would likely allow these to be viable. If these measurements are not possible, adapting the method provided to reduce the fitting region to something which is able to be validated would at least allow two more usable fits over small areas.

5.3 Shape detail

The model would be significantly improved if it were to take into account the finer detail in the insert cut-out shapes. To do this, it would be beneficial to have finer detail of the measurement data available. The first proposed method of improvement is to remove the deterministic lateral scatter effects discussed in Section 1.4.1. This would be done by using either the area, or approximate area (if area is unavailable), of the cut-out. The second is to make use of the sector integration method [27] discussed in Section 1.6.4. The third would be to make direct use of the parameters area and perimeter, instead of width and aspect ratio.

The deterministic lateral scatter effects are able to be predicted by finding the volume under a 3D Gaussian bounded by the insert cut-out shape. In order to define this Gaussian, its $\sigma$ is the electron spread at $d_{\text{max}}$. This could be approximated by stepping through various $\sigma$ in order to find the value which best fit the data. However, the methods discussed in Section 1.6.3 would be significantly more accurate in calculating this spread. Once this deterministic effect is removed, what remains would primarily be the cut-out shielding and collimator effects. These would be able to be fitted separately.

To implement the sector integration method, a range of circles is required. These circles could potentially be predicted using the modelling within this thesis. From this point these circles could then be used to predict cut-out factors, with significantly finer shape detail, via the sector integration method. This would allow for a fit that predicted cut-outs, with fine detail, and did not require measurements beyond those already conducted clinically.

Taking the perimeter/area relationship to cut-out factors discussed by Nair et al [3], these two parameters could be separated out and fitted in 3D in much the same way as the model used in this project. Potentially this could allow for greater determination of the shielding and cut-out collimator effects, and hence a better determination of cut-out factor with respect to varied shape.

5.4 Possibility for an adapted pencil beam model

A significant advantage of the pencil beam model over the theory described herein is its consideration of patient geometry. There is a possibility that an adapted pencil beam model could be produced using the methods described here. By comparing the measured output factors to a pencil beam prediction, given water tank geometry, the primary errors would be due to the unaccounted for shielding effects. This shielding effect factor can then be modelled using the methods described within this thesis. The final output factor can then be calculated using the pencil beam with patient geometry and then adjusted using the modelled shielding effects factor. It is predicted that this would provide an output factor that was 1-2% closer to the actual value.
5.5 Conclusion

This thesis discusses a method for the use of clinically collected measurements in the modelling and prediction of output factors for electron beams used in radiotherapy. This method reduces the frequency with which new factors must be measured, while providing quality assurance for newly measured factors. This was achieved using an adapted spline smoothing model to fit against the recorded widths and aspect ratios of the clinical cut-outs. Through computational resampling methods, it was found that five of the applicator/energy data set models were valid. The uncertainties in prediction were found to vary between 0.3% and 2%. This technique has the potential to stand alone as it is, or to be adapted/combined with previous methods such as the sector integration method or the pencil beam model. This provides a significant improvement for clinics who already interpolate output factors from previous measurements and will reduce the frequency with which measuring of new output factors is required.
Appendix A

Proof of parameter relationship

A.1 Ellipse

The area ($A$) of an ellipse is given in terms of width ($W$) and aspect ratio ($R$) is given in Equation A.1

$$A = \frac{\pi W^2}{4R} \quad (A.1)$$

The perimeter ($P$) of an ellipse can be approximated using Ramanujan’s formula with the radius along the major axis ($a$) and radius along the minor axis ($b$). This is given in Equation A.2:

$$P \approx \pi \left[ 3(a + b) - \sqrt{(3a + b)(a + 3b)} \right] \quad (A.2)$$

Converting Ramanujan’s formula to the use of width ($W$) and aspect ratio ($R$) the following is given in Equation A.3:

$$P \approx \frac{\pi W}{2R} \left[ 3(R + 1) - \sqrt{(3R + 1)(R + 3)} \right] \quad (A.3)$$

Then given that the first order Taylor expansion about $R = 1$, within reason for $0.5 < R \leq 1$, is given by Equation A.4

$$\left[ 3(R + 1) - \sqrt{(3R + 1)(R + 3)} \right] \approx R + 1 \quad (A.4)$$

The approximate perimeter formula given in Equation A.3 can be reduced to that given in Equation A.5:

$$P \approx \pi W \cdot \frac{R + 1}{2R} \quad (A.5)$$
Therefore, using Equations A.1 and A.5 the perimeter/area ratio can be calculated to be that given by Equation A.6:

\[
\frac{P}{A} \approx \frac{2(R + 1)}{W} \quad (A.6)
\]

### A.2 Rectangle

The perimeter of a rectangle in terms of width and aspect ratio is given by Equation A.7:

\[
P = 2W \cdot \frac{R + 1}{R} \quad (A.7)
\]

The area of a rectangle in terms of width and aspect ratio is given by Equation A.8:

\[
A = \frac{W^2}{R} \quad (A.8)
\]

Therefore the perimeter/area aspect ratio is given by Equation A.9 which is the same as Equation A.6 given above.

\[
\frac{P}{A} = \frac{2(R + 1)}{W} \quad (A.9)
\]
Appendix B

Primary code used

B.1 Program and package requirements

To run this code Octave and the octave-forge splines package is required. Octave can be acquired via pre-built installers at http://octave.sourceforge.net/ or by using your favourite package manager within Linux:

   sudo apt-get install octave

   The splines package may be downloaded from http://octave.sourceforge.net/splines/. To install the splines package run the following within Octave while in the working directory of the downloaded file:

   pkg install splines-1.2.7.tar.gz -auto

   Contact simon@biggs.email if you desire the full package of code required to run this model.

B.2 Variable explanations

The variables \texttt{wdth}, \texttt{ratio}, and \texttt{ins} are vectors containing the list of measured insert cut-out widths, ratios, and measured factors. When \texttt{app} is used this refers to the applicator size. Variables \texttt{x} and \texttt{y} are consistently used to represent the particular width and ratio in question.
B.3 The fit

B.3.1 Point prediction

```matlab
function prediction = point_prediction(x,y,wdth,ratio,ins)

% Defining the spline parameter p, and the weighting
% parameter sigma
p = 0.7; sigma = 0.1;

% Weight each of the points according to their distance in
% the ratio direction
weighting = normpdf(ratio,y,sigma);

% Use a smoothing spline in order to predict the points
prediction = smoothing_spline(wdth,ins,p,x,weighting);
```
B.3.2 Smoothing spline

function fit = smoothing_spline(x,y,p,xi,w)

% This function is adapted from "dedup.m" found within the
% splines package on octave forge. The purpose of this
% function is to remove duplicate x values in order to
% prepare for the "csaps" function.

[x,i] = sort(x);
y = y(i);
w = w(i);

h = diff(x);

if any(h <= 0)
    hh = ones(size(x));
    hh(2:end) = cumsum(h > 0) + 1;

    wnew = accumarray(hh', w);
    x = accumarray(hh', x .* w) ./ wnew;
    y = accumarray(hh', y .* w) ./ wnew;
    w = wnew;

    x = x'; y = y'; w = w';
end

% Run the cubic smoothing spline function found within the
% splines package on octave forge
fit = csaps(x,y,p,xi,w);
B.4 The region parameters

B.4.1 Angle gap

function gap = angle_gap(app,x,y,wdth,ratio)

% Scaling of the widths to be on the same order as the
% aspect ratio
wdth = wdth/app;
x = x/app;

% Finding the axis lengths between the point in question
% and all the other points
dx = wdth-x; dy = ratio-y;

% If a point in question is exactly equal to another point
% return 0 gap.
if any(dx==0 & dy==0)
    gap = 0;
    return
end

% Calculate the angle between the data to the point in
% question and the x-axis.
theta = atan(dy./dx);
theta(dx<0) = theta(dx<0) + pi;
theta(dx>0 & dy<0) = theta(dx>0 & dy<0) + 2*pi;
theta(dx==0 & dy>0) = pi/2;
theta(dx==0 & dy<0) = 3*pi/2;

% Sort the angles in ascending order
test = sort(theta);
% Copy the first angle to the final position also so that
% there is a cyclic boundary
test = [test,test(1) + 2*pi];
% Take the difference between all adjacent angles. Find the
% maximum of these and convert into degrees.
gap = max(diff(test))*180/pi;
B.4.2 Fit give

function give = fit_give(x,y,wdth,ratio,ins)

% Use the fit to find the predicted factor for the given
% width and ratio
inTrue = point_prediction(x,y,wdth,ratio,ins);

% Expand the data sets to include the width and ratio in
% question as a new data point
wdAdj = [wdth,x];
rtAdj = [ratio,y];
% Create an extreme positive point and an extreme negative
% point
inPos = [ins,inTrue + 1];
inNeg = [ins,inTrue - 1];

% Using the new data sets created, containing the extreme
% points, predict the new cut-out factor.
calcPos = point_prediction(x,y,wdAdj,rtAdj,inPos);
calcNeg = point_prediction(x,y,wdAdj,rtAdj,inNeg);

% Calculate the positive and negative give
givePos = (calcPos - inTrue);
giveNeg = (inTrue - calcNeg);

% Return the maximum value
give = max([givePos, giveNeg]);
B.4.3 Fit slope

function slope = fit_slope(app,x,y,wdth,ratio,ins)

% Defining a numerical dx and dy
dx = range(width)/1000; dy = range(ratio)/1000;

% Calculating the partial derivative in the direction of
% width
xPartial = ( ...  
    point_prediction(x+dx/2,y,wdth,ratio,ins) - ...  
    point_prediction(x-dx/2,y,wdth,ratio,ins) ...  
) / dx;

% Calculating the partial derivative in the direction of
% ratio. This includes the applicator adjustment.
yPartial = ( ...  
    point_prediction(x,y+dy/2,wdth,ratio,ins) - ...  
    point_prediction(x,y-dy/2,wdth,ratio,ins) ...  
) / (dy * app);

% Calculating the final slope
slope = sqrt(xPartial^2 + yPartial^2);
Bibliography


