

UNIVERSITY OF ADELAIDE



THE UNIVERSITY
of ADELAIDE
DOCTORAL THESIS

**A radiobiological Markov model for
aiding decision making in proton
therapy referral**

ANNABELLE MARY AUSTIN

Supervisors: Dr. Scott Penfold, Dr. Michael Douglass
& Dr. Giang Nguyen

*A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy*

in the

School of Physical Sciences
University of Adelaide

September 2019

Declaration of Authorship

- I, Annabelle Mary Austin, certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.
- I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.
- I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.
- I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Signed:

Date: 19/09/19

Peer Reviewed Publications

Published

- P1. Austin, A.M., Douglass, M.J.J., Nguyen, G.T. & Penfold, S.N. A radiobiological Markov simulation tool for aiding decision making in proton therapy referral. *Physica Medica*. 2017; 44:72–82.

Submitted for Publication

- P2. Austin, A.M., Douglass, M.J.J., Nguyen, G.T. & Penfold, S.N. Patient selection for proton therapy: A radiobiological fuzzy Markov model incorporating robust plan analysis. *Mathematical Medicine and Biology*. 2019 (Submitted May 2019).
- P3. Austin, A.M., Douglass, M.J.J., Nguyen, G.T., Daltsen, R., Le, H., Gorayski, P., Tee, H., Penniment, M. & Penfold, S.N.. Cost-effectiveness of proton therapy in treating base of skull chordoma. *Australasian Physical and Engineering Sciences in Medicine*. 2019 (Submitted May 2019).
- P4. Austin, A.M., Douglass, M.J.J., Nguyen, G.T., Cunningham, L., Le, H., Hu, Y. & Penfold, S.N.. Individualized selection of left-sided breast cancer patients for proton therapy based on cost-effectiveness. *International Journal of Particle Therapy*. 2019 (Submitted May 2019).

These publications are included within this thesis. When referred to in the text the reference number is prefixed by a ‘P’. For example, the first publication in this list is referred to as P1.

Grants and Scholarships

- Recipient of a Cancer Council (South Australia branch) Beat Cancer Project travel grant.
- Australian Postgraduate Award (APA) 2016 - 2019.

Conference Presentations and Awards

Local

- Austin, A.M., Douglass, M.J.J., Nguyen, G.T. & Penfold, S.N.. *A radiobiological Markov simulation tool for aiding decision making in proton therapy referral*. ACPSEM (SA/NT) Student Paper Night, 2016, **4th Place**.
- Austin, A.M., Douglass, M.J.J., Nguyen, G.T. & Penfold, S.N.. *A radiobiological Markov simulation tool for aiding decision making in proton therapy referral*. ACPSEM (SA/NT) Student Paper Night, 2017, **2nd Place**.
- Austin, A.M., Douglass, M.J.J., Nguyen, G.T. & Penfold, S.N.. *A radiobiological Markov simulation tool for aiding decision making in proton therapy referral*. ACPSEM (SA/NT) Student Paper Night, 2018.
- Austin, A.M., Douglass, M.J.J., Nguyen, G.T. & Penfold, S.N.. *To proton or not to proton?*. Three minute thesis competition, 2016.

National

- Austin, A.M., Douglass, M.J.J., Nguyen, G.T. & Penfold, S.N.. *A radiobiological simulation tool for aiding decision making in proton therapy referral*. EPSM, 2017, Hobart. **Prize for best radiation therapy presentation**.

International

- Austin, A.M., Douglass, M.J.J., Nguyen, G.T. & Penfold, S.N.. *A radiobiological simulation tool for aiding decision making in proton therapy referral (Poster)*. ESTRO 37, Barcelona, Spain, 2018.

Abstract

Cancer is a highly prevalent disease that places a significant economic burden upon society. Radiotherapy is commonly utilised as a treatment for benign and malignant tumours. A fundamental challenge in radiotherapy is delivering a sufficient dose of radiation to eradicate a tumour while minimizing the dose deposited in surrounding healthy tissue. Excessive radiation damage to these tissues can result in treatment toxicities that may have adverse effects on patient quality of life.

Proton therapy offers the potential for increased sparing of normal tissue compared with X-ray therapy, which is more commonly used in radiotherapy. However, the degree of this sparing can be highly variable between patients. Furthermore, data from Phase III clinical trials can quickly become outdated due to the long follow-up times that are required to observe late effects, together with the rapid evolution of technology. The process of deciding whether to refer a patient for proton therapy can be complex as a result. In addition, proton therapy is significantly more expensive as a treatment compared with X-ray therapy. This suggests that patients who are expected to receive the greatest benefit should be prioritised. Computer models can offer a possible solution to this dilemma, by predicting the clinical outcome that may be expected as a result of a given treatment.

In this work, a Markov simulation tool was developed which is capable of producing such predictions and comparing proton and X-ray radiotherapy treatment plans on an individual patient basis. The radiobiological effect of a given treatment plan is estimated in terms of the probabilities of tumour control, radiation-induced injuries and radiation-induced second cancers. These are combined in the Markov model to efficiently estimate the clinical outcome resulting from a given treatment plan. This outcome is quantified in terms of the quality adjusted life expectancy (QALE), or number of quality adjusted life years (QALYs), which is an adjustment of the raw life expectancy to account for the effect of time spent with injury or disease. The result is a model that uses several input parameters to produce a single quantitative output, indicative of the relative quality of a treatment plan.

The predictions of the model can be affected by uncertainties in the radiobiological model parameters and uncertainties in dose delivery. The latter can arise as a result of changes in the target volume relative to the radiation field over the course of treatment. A consideration of these effects was incorporated into the model, as they have the potential to influence whether a patient is selected to receive proton therapy.

The cost-effectiveness of a treatment is of particular importance in the current resource limited healthcare environment. The Markov model was developed to include treatment costs, including treatment of radiation therapy side effects. An application of the model to a cohort of base of skull chordoma patients revealed that all patients could be treated with proton therapy cost-effectively due to the potential for sparing of critical structures. Base of skull chordoma is typically regarded as a standard indication for proton therapy. In contrast, in a study of a cohort of left-sided breast cancer patients, it was found that the majority of patients could not be treated cost-effectively with proton therapy. This was likely due to the cardiac toxicity rate being particularly low with the deep inspiration breath hold X-ray treatment technique used for the patients in this cohort, resulting in no significant advantage from proton therapy.

The developed model has the potential to form the basis of a clinically viable patient selection tool. However, the model requires external validation before being suitable for clinical implementation. Due to the limited availability of proton therapy, such a model may prove useful as Australia prepares to begin treating cancer patients with proton therapy.

Acknowledgements

The substantial contributions of many have allowed me to write this thesis. I may not be able to repay them for their efforts entirely, but I wish to at least acknowledge and thank them here.

I've been fortunate to receive high quality supervision from Doctor Scott Penfold, Doctor Michael Douglass and Doctor Giang Nguyen for the duration of my PhD. I've always been impressed and inspired by your high level of experience despite your young ages. For me, the past years have been an incredibly valuable learning experience. This has been enabled by your insightful advice and dedication to my project. I've admired how you have not allowed long-term overseas commitments to be a barrier to high quality supervision. To Giang in particular, I am grateful for the time you have taken to explain areas of mathematics I was unfamiliar with, and also for your patience in this regard. Thank you all for your enduring encouragement and for helping me to grow as a researcher.

To Doctor Alex Santos: While you're not a formal supervisor, you have been more than willing to provide me with useful advice and discussion on a number of occasions. Your expertise on second malignancy induction was particularly valuable to me.

I've enjoyed meeting the students who have come in and out of the office over the years. A special thanks goes to James Keal who often shared his programming expertise. You didn't necessarily give me the answers but rather assisted me in formulating the right questions to ask. Your input has been highly valuable to me in developing my programming skills. I've particularly enjoyed this process and I'd never imagined that I would become as proficient as I now am in this skill. I look forward to developing these skills for the duration of my career.

I am grateful for the financial support I have received from the Australian Government in the form of a Research Training Program Scholarship (formally Australian Postgraduate Award). While I acknowledge that the government faces many financial pressures, I hope that they will continue to support early career researchers and research more generally, as the benefits to society are worthwhile.

I was also fortunate to have received a Cancer Council SA Beat Cancer travel grant which gave me the opportunity to travel overseas, to experience a working proton therapy centre and to understand how patient selection tools may be implemented clinically.

To Anthea, my mum, I'm grateful for your ability to provide advice from afar, but also that you've shown me that a woman can succeed in a male dominated industry. I've always been amazed by the number of jobs you've been able to work in a day. Thank you for teaching me how to make every minute of the day count.

Ben, thank you for reminding me of the world outside my research, but also for being a cactus when the thesis called. Thank you for being overtly proud and refusing to allow me to forget the magnitude and importance of what I've been able to achieve.

To those I've met in life who know what it is like to suffer a debilitating medical condition. Thank you for inspiring me to walk into the office every day. It is all for you.

Contents

Declaration of Authorship	i
Peer Reviewed Publications	ii
Grants and Scholarships	iii
Conferences	iv
Abstract	v
Acknowledgements	vii
List of Figures	xiii
List of Tables	xiv
Abbreviations	xv
1 Introduction	1
1.1 Radiation Therapy	1
1.2 Proton therapy	2
1.2.1 Physics of proton therapy	3
1.2.2 Radiobiology	5
1.2.3 Availability and cost	6
1.3 Clinical trials of proton therapy	6
1.4 Model-based patient selection for proton therapy	8
1.5 Motivation and thesis structure	9
1.5.1 Motivation	9
1.5.2 Aims of current work	10
1.5.3 Thesis outline	11
2 Background and Literature Review	13

2.1	Decision making in radiation oncology	13
2.1.1	Patient selection strategies for proton therapy	14
2.1.2	Cost-effectiveness of proton therapy	18
2.1.3	Summary	20
2.2	Markov models	21
2.2.1	Theoretical background	21
2.2.2	Markov models in radiotherapy	22
2.2.2.1	Comparison with other prognosis modelling approaches	25
2.2.3	Quality of life concept and utilities	26
2.2.4	Summary	27
2.3	Radiobiological models	27
2.3.1	TCP	28
2.3.2	NTCP modelling	29
2.3.2.1	LKB model	29
2.3.2.2	Other NTCP models	30
2.3.3	Second cancer risk	31
2.3.4	Summary	33
3	Monte Carlo Evaluation	34
3.1	Introduction and motivation	34
3.2	Statement of contribution	35
3.2.1	Conception	35
3.2.2	Realisation	35
3.2.3	Documentation	35
3.3	Markov model details	49
3.3.1	Determining the number of Markov states	49
3.3.2	Transition probability calculation	50
3.3.2.1	Normal tissue complication probability (NTCP)	50
3.3.2.2	Second primary cancer induction probability (SPCIP)	51
3.3.2.3	Cancer and injury death	51
3.3.2.4	Scaling of death probabilities	52
3.3.3	The transition matrix	54
3.3.3.1	Representing states	55
3.3.3.2	Calculation of transition probabilities	56
3.3.4	Monte Carlo simulation	59
3.3.5	Analysis of model output	60
3.4	Model verification	61
3.4.1	Zero dose, tumour absent case	61
3.4.2	Zero dose, tumour present case	62
3.4.3	Uniform dose, tumour absent case	63
3.4.4	Statistical significance	65
3.5	Discussion	66
3.6	Conclusion	67

4	Analytic Evaluation	68
4.1	Introduction	68
4.2	Description of the AE model	69
4.2.1	Cancer staging	70
4.2.2	Inclusion of higher injury grades	71
4.2.3	Grade/stage dependence of death probabilities	74
4.2.4	Removal of negligible probabilities	75
4.2.5	Combining individual organ SPCIPs	75
4.2.6	Number of Markov states	77
4.3	Evaluation	78
4.3.1	Preventing redevelopment of injuries	78
4.3.2	Representing states	79
4.3.3	Structure of the transition matrix	81
4.3.4	Calculation of transition probabilities	82
4.4	Computational efficiency methods	84
4.4.1	Method of matrix condensation	85
4.4.2	Filling of the transition matrix in blocks	85
4.5	Model output	87
4.5.1	Raw life expectancy	87
4.5.2	Quality-adjusted life expectancy	89
4.5.3	Other metrics	90
4.6	Comparison with the MCE model	90
4.7	Discussion and conclusion	95
5	Model Uncertainties Incorporating Robust Plan Analysis	97
5.1	Introduction and motivation	97
5.2	Statement of contribution	98
5.2.1	Conception	98
5.2.2	Realisation	98
5.2.3	Documentation	98
5.3	Fuzzy Markov model details	132
5.3.1	NTCP calculation	132
5.3.2	Interval arithmetic	133
5.3.3	Quality of life utility weights	134
5.3.4	Transition matrix	135
5.4	Discussion and conclusion	135
6	Cost-effectiveness Model	136
6.1	Introduction and motivation	136
6.2	Statement of contribution	137
6.2.1	Conception	137
6.2.2	Realisation	137
6.2.3	Documentation	137
6.3	Discussion and conclusion	161

7	Cost-effectiveness of Proton Therapy: Breast Cancer	162
7.1	Introduction and motivation	162
7.2	Statement of contribution	163
7.2.1	Conception	163
7.2.2	Realisation	163
7.2.3	Documentation	164
7.3	Calculation of transition probabilities from DVH data	186
7.3.1	Tumour control probability (TCP) calculation	186
7.3.2	Normal tissue complication probabilities	186
7.3.2.1	Pneumonitis	186
7.3.2.2	Heart disease	187
7.3.3	Second primary cancer induction probabilities	188
7.4	Estimation of costs	189
7.4.1	Radiation therapy	189
7.4.2	Cancer	190
7.4.3	Heart disease	190
7.4.4	Pneumonitis	190
7.5	Dose-dependent transition probabilities by patient	191
7.6	Mean organ doses by patient	192
7.7	Discussion and conclusion	193
8	Conclusions and Future Work	195
8.1	Conclusions	195
8.2	Future work	197
	Bibliography	199

List of Figures

1.1	Proton beam delivery systems	4
1.2	Spread out Bragg peak	5
2.1	Difference in NTCP between protons and photons for a given dose	15
2.2	Patient transitions between states	24
2.3	State transition diagram	24
2.4	Decision tree example	25
3.1	Demonstration of the conversion of the transition probabilities from absolute to cumulative	59
3.2	Survival curve for Monte Carlo simulation verification (unrelated death) .	62
3.3	Survival curve for Monte Carlo simulation verification (cancer death) . .	63
3.4	Injured fraction (Monte Carlo simulation verification)	64
3.5	Survival curve for Monte Carlo simulation verification (injury death) . .	65
4.1	AE model verification (survival and cancers)	73
4.2	AE model verification (injuries)	74
4.3	Comparison of output of the MCE model and the AE model (survival) .	92
4.4	Comparison of output of the MCE model and the AE model (injury incidence)	93
4.5	Comparison of output of the MCE model and the AE model (second cancer incidence)	93
4.6	Computation time comparison	95

List of Tables

3.1	Log-rank test results for the publication P1	61
3.2	Chi-squared test results for Monte Carlo simulation verification	66
4.1	Comparison of output of the MCE model and the AE model	92
4.2	Computation time comparison	94
5.1	Parameters used for the calculation of the NTCP for each injury using the LKB model	133
7.1	Parameters used for the calculation of the SPCIP for each year after treatment. Data was not provided for the heart.	189
7.2	Transition probabilities calculated based on the dose to each patient. Duplicates are indicated by an asterisk.	192
7.3	Mean doses to the heart and lung for each patient and treatment modality. Duplicates are indicated by an asterisk.	193

Abbreviations

AEM	A nalytically E valuated M odel
BOSCh	B ase O f S kull C hordoma
CT	C omputer T omography
CTV	C linical T arget V olume
DIBH	D eep I nspiration B reath H old
DVH	D ose V olume H istogram
ICER	I ncremental C ost E ffectiveness R atio
IMRT	I ntensity M odulated R adiotherapy
IMPT	I ntensity M odulated P roton T herapy
LET	L inear E nergy T ransfer
LQ	L inear Q uadratic
MC	M onte C arlo
MCEM	M onte C arlo E valuated M odel
NTCP	N ormal T issue C omplication P robability
OAR	O rgan A t risk
PTV	P lanning T arget V olume
QALE	Q uality A justed L ife E xpectancy
QALY	Q uality A justed L ife Y ear
RBE	R elative B iological E ffectiveness
SPCIP	S econd P imary C ancer I nduction P robability
TCP	T umour C ontrol P robability
TPS	T reatment P lanning S ystem
VMAT	V olumetric M odulated A rc T herapy

Chapter 1

Introduction

In 2011, cancer was the leading cause of disease burden in Australia [1]. In 2017 it was estimated that more than 130,000 Australians were diagnosed with cancer, corresponding to an average of 367 diagnoses per day [1]. It is likely that this will increase given the ageing population. Therefore, it is important to ensure access to high quality treatments. There are several treatment options available for cancer, including surgery, radiation therapy, chemotherapy and immunotherapy.

1.1 Radiation Therapy

Radiation therapy utilises ionising radiation to damage the DNA in tumour cells, hence facilitating the eradication of the tumour. High energy X-rays are typically used for this purpose. Radiotherapy treatments can be broadly classified as either external beam radiotherapy or brachytherapy [2]. These differ in the location of the source of radiation with respect to the target within the patient. External beam radiotherapy involves a source at a distance from the patient. Brachytherapy involves placing the source within the patient, or immediately abutting the patient for skin treatments. It has been estimated that 48% of Australian cancer patients should receive external beam radiation therapy at least once during their treatment [3].

The primary goal of radiotherapy is to deliver a maximum dose of radiation to the tumour volume, while minimising the dose to surrounding healthy tissue. A major limiting factor on the maximum dose that can be delivered to the tumour, and consequently the chance of a successful treatment, is the tolerance of normal tissue. Depending on the magnitude and location of the radiation dose received by normal tissue, any damage may be sufficient to result in treatment side effects. These can have a significant impact on the quality of life of the patient. Therefore, many techniques are employed in radiotherapy clinics to minimise the amount of radiation received by normal tissue. This is a rapidly evolving area of research. For example, newer treatment technologies such as intensity modulated radiotherapy (IMRT) and volumetric modulated arc therapy (VMAT) allow an escalated dose to be delivered to the tumour volume without increasing the dose received by surrounding healthy tissue. The use of charged particles, such as protons or carbon ions, rather than photons can also yield superior dose distributions in certain cases.

Prior to receiving a course of radiotherapy, a computed tomography (CT) scan of the patient is normally used by a treatment planner to develop a treatment plan. While the primary purpose of this is to determine optimal treatment strategy, it also allows estimates of the dose that would be deposited in the tumour and in normal tissue to be obtained.

1.2 Proton therapy

With technological advances, IMRT and VMAT have become the most common treatment modalities in radiation oncology. However, the use of proton therapy has been becoming increasingly available in recent years. Proton therapy has the advantage of being able to significantly reduce the radiation dose that is deposited in tissues outside the tumour volume. Protons have been used for therapeutic purposes since the mid-1950s [4].

1.2.1 Physics of proton therapy

A description of the basic physics of proton therapy is given by Khan [5]. Proton therapy involves a beam of high energy protons that have been accelerated using either a cyclotron or synchrotron. The more favourable dose distribution of proton therapy compared with photons can be explained by the difference in the way in which protons interact with matter. The rate of energy loss of a particle per unit path length in a medium is represented by the stopping power. The dose absorbed in the medium is related to this energy loss. The stopping power increases as the particle velocity decreases, with the energy loss reaching a maximum near the end of the particle range. For protons, the sharp increase in the deposited dose near the end of the particle range is known as the Bragg peak (Figure 1.2).

The characteristically narrow Bragg peak is not clinically useful for irradiating tumours which typically have a much larger physical extent. A spread-out Bragg peak (SOBP) can be obtained through the superposition of several beams of different energies, as depicted in Figure 1.2. The energy range is chosen such that the range of the particles in the highest energy beam is sufficient to reach the distal edge of the target volume, and this is superimposed on beams of decreasing energy. The result is a more uniform dose over the tumour region compared with a single beam.

A beam delivery system is necessary to spread the narrow beam to the required field size for treatment. This system can be either of two types: passive scattering or pencil beam scanning, which are described by Khan [6]. Passive scattering is where the beam is scattered using thin foils. For these systems, range modulator wheels are utilised to obtain a range of energies resulting in the SOBP, however the longitudinal width of the SOBP is also constant across the field which results in larger doses outside of the target volume (see Figure 1.1). Range compensators are required for the isodose lines to conform to the distal edge of the target volume. The range compensators and modulators increase the number of interactions that the beam undergoes outside of the patient, which increases the neutron contamination.

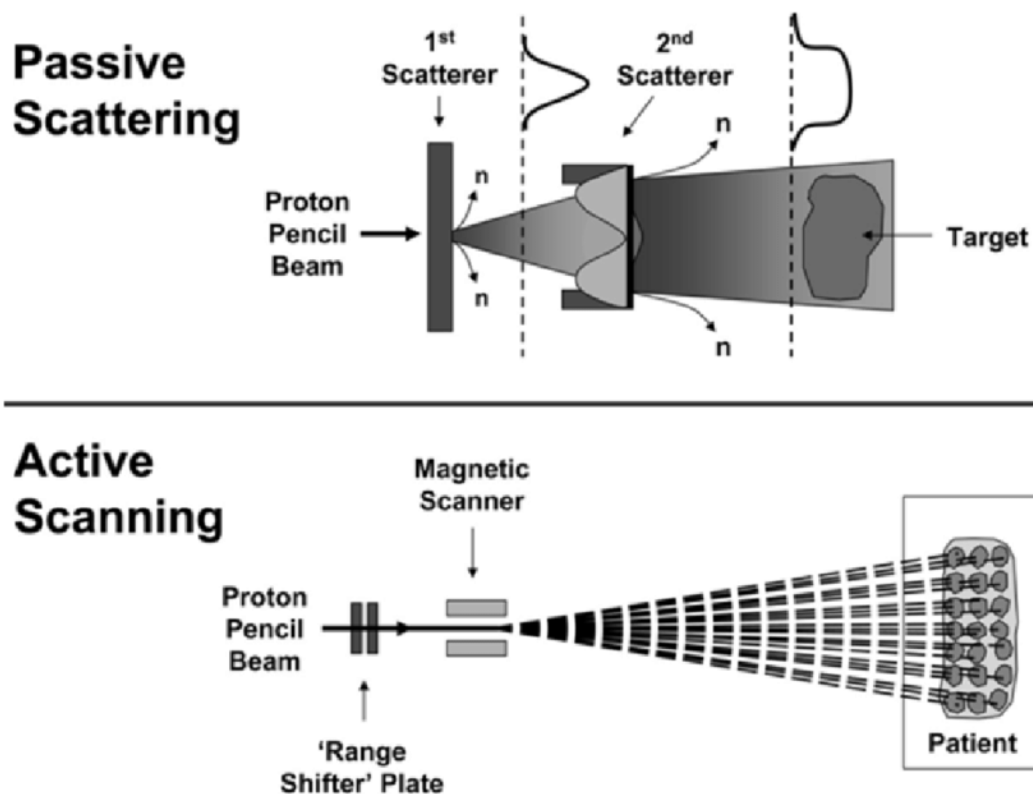


FIGURE 1.1: Comparison of passive scattering and active (pencil beam) scanning beam delivery systems. Passive scattering systems result in larger doses outside of the target volume. Source: [7].

Scanning involves scanning the thin pencil beam over the target volume with magnets. In this system, range modulation is achieved with energy degraders (objects of variable thickness which are placed in the path of the beam) at the exit of the accelerator or using variable energy accelerators. The use of pencil beam scanning is becoming more common, due to the superior dose distributions achievable compared to passive scattering [8]. In addition, the flexibility of scanning makes it an ideal technique for the delivery of intensity modulated proton therapy (IMPT). However, a disadvantage is the greater sensitivity of scanning to organ motion due to respiration. This effect, known as the interplay effect, can result in the delivery of a suboptimal dose. Techniques have been developed to reduce this effect, such as the isolayered rescanning approach of Karder *et al.* [9] which involves simulating the effect of motion so it can be accounted for.

Figure 1.2 illustrates the superior dose distribution of a proton beam with a SOBP formed by several proton beams, compared with a single X-ray beam. In contrast to protons,

photons are attenuated exponentially in matter resulting in an exponential decrease in deposited dose with depth. The physical dose deposition properties of protons result in a conformal tumour dose and close to zero exit dose compared with photons. This makes it possible to increase tumour dose without increasing dose to normal tissue.

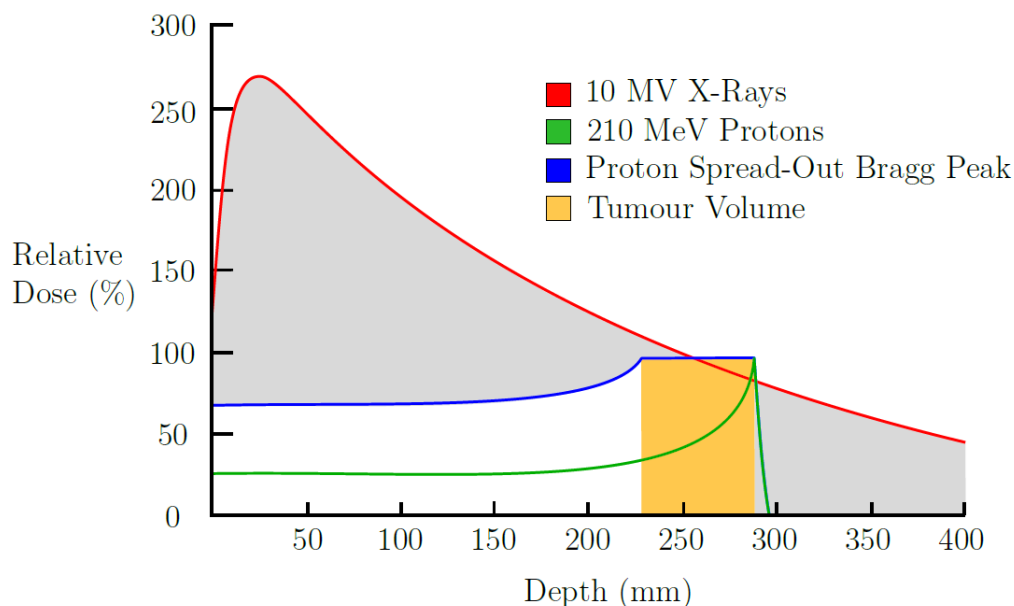


FIGURE 1.2: Percentage depth dose (PDD) for a 210 MeV proton beam (green), and several proton beams of different energies resulting in a SOBP (blue). The dose deposition for a photon beam is also shown for comparison (red). Reproduced with permission from Michael Douglass.

1.2.2 Radiobiology

The relative biological effectiveness (RBE) of any radiation is the ratio of the dose from 250 kV X-rays required to produce a given biological effect to the dose from the radiation of interest to produce the same biological effect. While there is uncertainty regarding the RBE of protons, most clinics adopt a value of 1.1 [10]. Therefore, a 10% greater physical dose from photons is required to produce the same biological effect as for protons. This value has been determined through extensive studies and serves only as a rough approximation for all tissues. It can also vary with the depth of the protons, with higher values in the distal part of the SOBP [11]. The uncertainty in the RBE is particularly

problematic considering magnitude of the end of range dose [12]. A misplaced Bragg peak could result in considerable dose to healthy tissue.

Secondary radiation arising from nuclear interactions both inside and outside of the patient is also an important consideration in the biological effect of protons [13]. For example, neutrons that are produced in these interactions can carry energy a significant distance from the interaction site. These neutrons have a high RBE and neutrons generated outside of the patient increase the integral dose.

1.2.3 Availability and cost

While proton therapy can potentially offer a superior dose distribution compared with X-ray therapy, it is also a more expensive form of treatment with limited availability. This is particularly true in Australia, where there are currently no proton therapy treatment facilities and patients must be sent overseas for proton therapy (although as of 2018 several facilities were in the early stages of planning).

The cost of proton therapy is typically estimated to be approximately 2.5 times that of X-ray therapy [14]. However, it is likely that this ratio will decrease over time. Goitein and Jermann [15] estimated that the ratio could decrease to 1.7. In some situations, it may be possible to justify the expense of proton therapy if it is anticipated that it will lead to a lower cost of treating radiation-induced complications than photon radiotherapy.

1.3 Clinical trials of proton therapy

Clinical trials are an important component of research of cancer treatments. An introduction to the concept of clinical trials of ion beam therapy is given by Cox [16]. Clinical trials in radiation oncology can be classified as follows:

- Phase I: the safety of a new treatment is tested on a small cohort of patients.
- Phase II: the efficacy of the treatment is tested on a larger cohort of patients.

- Phase III: randomised control trials (RCTs). The efficacy of the newer treatment is compared with a standard treatment. These trials are considered the “gold standard” as valid conclusions can be drawn.

While proton therapy can result in a more favourable dose distribution in a patient compared with X-ray therapy, there is uncertainty as to whether this distribution corresponds to a significant difference in clinical outcome. There is evidence of improved quality of life and patient reported outcomes for selected cancers after receiving proton treatment [17]. However, at the end of 2018 no Phase III clinical trials were able to show the benefit of protons. However, a number were being conducted that are detailed by Mishra *et al.* [18].

Concern has been raised regarding the lack of positive Phase III clinical trials of proton therapy, recommending that they should be required before proton therapy is adopted as standard treatment [19–22]. However, others have argued that positive Phase III clinical trials of proton therapy should not be a requirement [23, 24]. For example, Suit *et al.* [23] note that if it were not for the larger cost, then proton therapy would be adopted as a standard treatment. It has been well established that protons can deliver an equal or more favourable dose distribution compared with X-rays and hence the only disadvantage of proton therapy is treatment cost. However, Suit *et al.* [23] and Goitein and Cox [24] argue that cost-effectiveness alone is insufficient grounds to justify the need for an RCT, especially when the expensive treatment will also likely produce the better outcome. Furthermore, in order for a clinical trial to be morally justifiable, there must be uncertainty regarding the relative benefit of the two arms of the trial [25], also known as equipoise. An equivalent or superior dose distribution can be expected from proton therapy compared to X-ray therapy, and hence it is not possible to ethically assign patients to arms of a trial. In the opinion of Suit *et al.* [23], resources would be better allocated to improving proton therapy treatment.

Dahl [26] has also noted that in the past other technologies in radiation oncology such as cobalt units and IMRT have not required positive Phase III clinical trials to be adopted as part of routine clinical practice. Instead, cobalt units were adopted based on improved modelled dose distributions.

There are other reasons why RCTs may not be suitable for investigating and quantifying the relative clinical benefit of protons. A significant challenge for Phase III clinical trials is that the treatment techniques being compared evolve rapidly and can change significantly over the time period required to conduct clinical trials. Furthermore, some radiation-induced injuries take several years to develop. For example, heart disease often occurs 5 years after treatment and can occur up to 20 years later [27]. The development of second malignancies can have a greater latency period. It is not practical to conduct clinical trials over time periods of this magnitude as once the results are obtained, they will not be relevant to the current technology.

Enrolment of patients in these trials is another issue, particularly if health insurance is a limiting factor. Study biases can result if selected groups of patients are not enrolled in the proton arm of the trial due to lack of insurance coverage. In addition, Glimelius and Montelius [22] noted that it is possible that proton therapy can benefit more rare types of cancers which would make it difficult to recruit sufficient members for a trial in a reasonable time frame.

It is apparent that Phase III clinical trials may not be appropriate for determining whether the superior dose distribution of proton therapy translates to a clinical benefit. Furthermore, the magnitude of any clinical benefit is likely to be highly variable between patients. In addition, the limited availability of the treatment suggests that it should be prescribed preferentially to patients who are expected to receive the greatest benefit compared with X-ray therapy. The result is oncologists being presented with a challenge when deciding whether to refer a patient for proton therapy.

1.4 Model-based patient selection for proton therapy

As discussed in Section 1.3, deciding which patients to treat based on the results of Phase III clinical trials of proton therapy may not be the best approach. An alternative approach is to use evidence-based mathematical models to provide predictions of clinical outcomes. The concept of model-based patient selection was proposed by Langendijk *et al.* [28]. The approach aims to select patients to receive proton therapy based on the

predicted clinical outcome relative to the clinical outcome of X-ray therapy. Specifically, the clinical benefit is estimated in terms of the normal tissue complication probability (NTCP), that is, the probability of radiation-induced injury. The method was proposed as a means of introducing proton therapy into the Dutch health care system, in lieu of evidence in the form of RCTs. The model is discussed in more detail in Section 2.1.1. The limitations of the approach are that there is no consideration of treatment failure or of radiation induced secondary malignancies.

1.5 Motivation and thesis structure

1.5.1 Motivation

Compared with X-ray therapy, proton therapy has the potential to reduce radiation treatment complication rates in selected groups of patients. However, it is also significantly more expensive and has limited availability. Therefore, it is important to identify patients who are likely to experience an improved clinical outcome if treated with proton therapy. For the reasons discussed in Section 1.3, randomised clinical trials have not been conducted until recently. This leads to a more challenging treatment referral process. Modelling studies may provide an alternative approach to assessing proton therapy outcomes. As personalized medicine becomes increasingly important, it is advantageous to quantify the clinical benefit of proton therapy compared with X-ray therapy on an individual patient basis.

A proton therapy patient selection model has been developed [28] to address the issue of the limited availability of the treatment in the Netherlands where it has recently become available. However, there remains potential for alternative, more refined approaches. The model does not consider the possibility of treatment failure or the induction of radiation-induced malignancies. The latter is a particularly important consideration given the expected difference in integral dose between a proton and photon treatment. Furthermore, it is important to account for the effect of radiation-induced malignancies on the clinical

outcome of younger patients, as they have a greater remaining lifetime over which to develop such a cancer.

While proton therapy was unavailable in Australia at the time of writing (early 2019), construction of a treatment facility in Adelaide was close to commencement. Facilities in other cities were also in the early stages of planning. In the coming years, Australia will continue to face the issue of limited availability of this lifesaving treatment. An evidence-based proton therapy patient selection model can aid in this respect.

1.5.2 Aims of current work

The aim of the current work was to develop a toolkit with the ability to predict the clinical outcome of a given radiotherapy treatment on an individual patient basis. This prediction may then be used to assess the optimal radiotherapy treatment strategy for an individual patient.

The toolkit is based on a Markov model which is a type of stochastic model. These have been utilised in radiotherapy to predict patient outcomes and are described in more detail in Section 2.2. The model input is the dose-volume histogram (DVH) data corresponding to a particular patient's treatment plans, one for proton therapy and one for X-ray therapy. This data specifies the amount of radiation the tumour volume and each of the organs at risk receives. The patient age and gender are also necessary inputs.

For a given treatment plan, radiobiological models are used to determine the probabilities of tumour control, radiation-induced injuries and radiation-induced second cancers. These are combined in the Markov model into a single metric which is output by the model, the quality-adjusted life expectancy (QALE). This is an adjustment of the predicted life expectancy for the effect of treatment complications and is a more holistic representation of the treatment outcome than simply the NTCP. The difference in QALE between two treatment plans (one proton, one photon) provides a quantitative estimate of the relative benefit a patient can expect if treated with proton therapy.

The intended usage of the model is to select patients to receive proton therapy based on a comparison of predicted QALEs for their proton and X-ray radiotherapy plans. The

strength of the toolkit is its ability to combine several competing treatment factors (such as tumour control and complication risk) into a single metric that is indicative of the quality of a treatment plan. In addition, the toolkit is able to consider any treatment site and any number of radiation-induced injuries when determining the QALE. While it was developed to compare proton therapy with X-ray therapy, the toolkit has the flexibility to compare any two radiation treatment techniques provided a treatment plan can be provided for each.

In addition to developing the Markov toolkit to select patients to receive proton therapy, the aims of the current work were to:

1. Incorporate the effect of treatment and model parameter uncertainties on the output of the tool;
2. Incorporate an analysis of treatment cost-effectiveness into the tool;
3. Given a cohort of base of skull chordoma patients, determine whether this indication can be treated with proton therapy cost-effectively; and
4. Investigate which patients, if any, would be selected to receive proton therapy from a cohort of left-sided breast cancer patients.

1.5.3 Thesis outline

This thesis includes several publications, which form the basis of some chapters. Other chapters have a conventional format.

Chapter 2 provides a review of patient selection methods in proton therapy, Markov models and radiobiological models. Necessary background information is also provided.

Chapter 3 outlines the first model that was developed, which is based on a Monte Carlo simulation. The methodology of the development of the model is presented, along with a demonstration of the output with an example patient. The publication P1 forms the basis of this chapter.

Chapter 4 outlines the second model that was developed, where the solution is calculated analytically, thus providing the exact solution. This approach also results in a significantly reduced computation time, while producing results that are equivalent to the first model (for a sufficiently large number of Monte Carlo iterations).

Chapter 5 details the methodology of including treatment and model parameter uncertainties in the model predictions. These uncertainties can potentially impact whether a patient is selected for proton therapy. The magnitude of these uncertainties are assessed and the output is demonstrated with an example patient. The publication P2 forms the basis of this chapter.

Chapter 6 details the incorporation of treatment cost-effectiveness analysis into the tool. Cost is often used to justify the need for more expensive treatments. The updated model is applied in a retrospective study to a cohort of base of skull chordoma patients to determine the cost-effectiveness of treating this indication with proton therapy. This cancer type is generally regarded as a standard indication for proton therapy. The publication P3 forms the basis of this chapter.

Chapter 7 describes the application of the toolkit to a cohort of left-sided breast cancer patients who had proton plans created retrospectively. Breast cancer is not generally considered as a standard indication for proton therapy. The aim was to determine which patients, if any, would be selected for proton therapy. The publication P4 forms the basis of this chapter.

Chapter 8 provides a conclusion of the thesis with a summary of the outcomes of the research. Directions for future work are recommended.

Chapter 2

Background and Literature Review

2.1 Decision making in radiation oncology

As discussed in Section 1.3, deciding on the best course of treatment for a cancer patient can be challenging. This is particularly true due to the limitation of data on the late effects of newer radiotherapy treatments. In addition, there are a number of factors that must be considered including treatment availability, quality of life and patient preferences [29]. While it is possible to assess and compare radiotherapy treatment plans visually, it is difficult to predict the degree to which a difference in dose distribution will affect the clinical outcome of a patient in terms of disease-free survival. Decision support systems based on mathematical models can assist in predicting the most likely patient outcome.

In a review of predictive models in radiation oncology, Lambin *et al.* [29] discussed the increasing importance of individualised medicine in the current healthcare landscape. They noted that an important challenge is quantitatively integrating clinical, imaging and molecular data, and that many current prediction models lack assessments of robustness, reproducibility and clinical utility.

An example of a clinical decision support tool was created by Brodin *et al.* [30] for radiotherapy plan comparison for Hodgkin lymphoma. Volumetric modulated arc therapy (VMAT) plans and 3D conformal plans were compared based on the risk of developing complications. As the disease can be spread through the body, normal tissue complication

probability (NTCP) models were included for multiple sites, weighted based on their reliability and applicability. Another decision aid has been developed by Smith *et al.* [31]. This tool was based on a Bayesian network and a Markov model and was used to compare different IMRT treatment plans for prostate cancer. In addition to complication risks, the tumour control probability (TCP) was also considered in this method.

Demand modelling can also inform clinical decision making in radiotherapy as well as assist in planning of radiotherapy services. An example was described by Delaney *et al.* [32], where the proportion of cancer cases eligible for radiotherapy was estimated using an optimal radiotherapy utilisation tree based on clinical guidelines and epidemiological data. It was found that in Australia, there were fewer patients who received radiotherapy than patients who would have received a benefit from the treatment.

2.1.1 Patient selection strategies for proton therapy

The importance of prioritising patients for proton therapy is discussed in Chapter 1. Ideally, those who will likely receive the most improved clinical outcome if treated with proton therapy compared with X-ray therapy should be selected. However, quantitatively defining clinical outcomes can be challenging. In a review of mathematical modelling for patient selection in proton therapy, Mee *et al.* [33] summarised a range of models that have been developed for this purpose. There are a variety of methods employed, including discrete event simulations, Markov models and NTCP-based models.

The concept of *in silico* clinical trials proposed by Langendijk *et al.* [28] has been introduced in Section 1.4. The aim of the approach is to provide evidence-based medicine in the absence of data from clinical trials. The approach involves generating a photon plan and a proton plan for a given patient. The dose to critical structures is calculated from each treatment plan and, together with validated NTCP models (see Section 2.3.2), is used to predict the likelihood of toxicity after each treatment. Hence, the difference in the dose distribution is translated to a difference in clinical outcome, as demonstrated in Figure 2.1 [28]. If the difference in the expected NTCP exceeds a defined threshold, then the patient is selected to receive proton therapy.

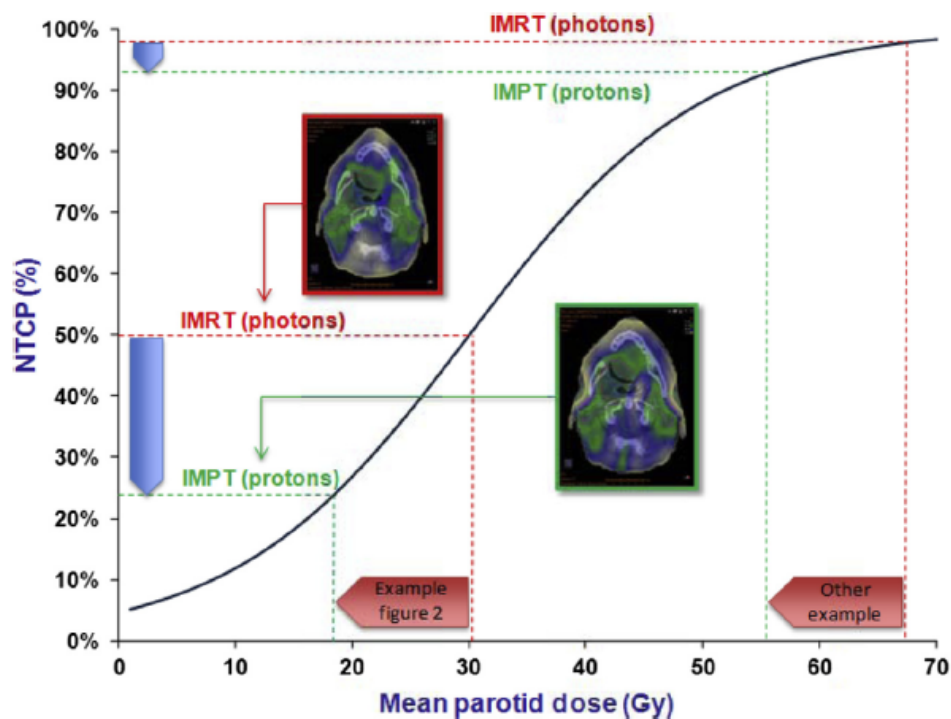


FIGURE 2.1: Demonstration of the translation of the difference in dose between a proton and photon treatment plan to a difference in normal tissue complication probability (NTCP), in this case xerostomia. Two examples are presented and it is apparent that the NTCP difference depends on both the dose difference and the dose region. Source: [28].

Proton therapy treatments began in the Netherlands in 2018 and this approach of patient selection has been adopted by Dutch health authorities. In addition, patient participation in follow-up programs is compulsory, which allows the compilation of valuable research data on patient outcomes after receiving proton therapy. This is particularly important to facilitate the validation and refinement of the NTCP models.

The question of whether or not proton therapy should be adopted with the absence of evidence from Phase III clinical trials has been discussed in Section 1.3. If it is adopted without such supporting evidence, then the risk is investing considerably in a treatment that may not improve clinical outcomes. Alternatively, if the introduction of proton therapy is delayed, then patients may receive sub-optimal treatment. Grutters *et al.* [34] analysed this trade-off using the method of real options analysis. This approach can assist in determining whether a treatment should be adopted, or whether a trial is required.

The optimal design of a trial can also be indicated in terms of sample size and follow-up time. For the case of proton therapy, the three options are to adopt the treatment without trial, adopt with a trial, or delay with a trial. Grutters *et al.* [34] have found the benefits of conducting trials to be sufficient to outweigh the costs when comparing proton therapy with stereotactic body radiotherapy in the treatment of stage I non-small cell lung cancer. Furthermore, they found that the expected net gain of adopting proton therapy is greater than that of not adopting it.

Langendijk *et al.* [28] and Widder *et al.* [35] also proposed that model-based approaches can be used for designing clinical trials of proton therapy, in addition to selecting patients to receive the treatment. Sequential prospective observation cohort studies are suggested as an alternative to traditional randomised control trials. The procedure of these cohort studies is as follows:

1. A *historical control group* is created. Each patient treated before protons were available (and therefore were treated with photons) has a proton plan created retrospectively. The difference in NTCP between each plan is calculated for each patient. The historical control group consists of the patients that would have been selected for proton therapy had it been available.
2. The *treatment group* consists of the patients selected to receive protons once the treatment is available using the same comparative planning procedure. The only difference is that the proton treatment plans are created before the treatment is delivered.
3. The outcomes of the two groups are compared using analysis procedures similar to those of traditional clinical trials.

The advantage of this approach is that the issue of ethically assigning patients to different arms of a clinical trial is removed (this issue is discussed in Section 1.3). All patients receive a proton treatment course if: (i) a proton treatment is available and (ii) they are expected to experience an improved clinical outcome compared with the photon treatment. Another advantage is that groups with similar characteristics are defined and biases are removed. If the groups were not defined carefully in this manner, then the treatment

group would be compared to all patients who did not receive protons. It is likely that many of these patients would not experience an enhanced clinical outcome from a proton treatment and this would skew the results. In this case, a positive result would mean that patients who would experience no benefit from receiving proton therapy would be selected to do so, while a negative result would mean that eligible patients would not be selected. However, the disadvantage of designing a clinical trial using this approach is that selection criteria may change over time, making it more difficult to compare current cohorts with historical cohorts.

The Proton Priority System developed by Bekelman *et al.* [36] aims to compare patients and select those who would likely receive the greatest benefit if treated with proton therapy. In this system, each patient is assigned a score which is a weighted sum of 7 domains. These include diagnosis, anatomic site, stage, co-morbidities, age, urgency, and clinical protocol. There are factors in each of these domains that are assigned a score between 0 and 10. Higher scores are assigned where a greater benefit from proton therapy is expected. For example, sites including the base of skull and spine are assigned 10 points, while the brain is assigned 5 points. Higher scores are also assigned to younger patients and more urgent cases. Proton therapy is more likely to benefit patients with localised cancers rather than metastatic cancer. The scores and weights were decided by the Proton Priority Oversight and Advisory Board established by the Roberts Proton Therapy Centre at the University of Pennsylvania. The board consists of oncologists, a nurse, a patient representative and a medical ethicist. The best available evidence and expert opinion are utilised in determining scores and weights. The framework also has the flexibility to evolve as new evidence and experience emerges.

An important principle of the Proton Priority System [36] is equity. Therefore, a patient's score should not be affected by sex, race, geography or insurance status. Bekelman *et al.* [36] investigated the association between score and receipt of proton therapy with insurance status, gender, race and geography. It was found that allocation depended on insurance status. Furthermore, the disadvantage of this approach is that it is more qualitative than quantitative compared with the approach proposed by Langendijk *et al.* [28].

As an alternative to individualised patient selection for proton therapy, Jakobi *et al.* [37] attempted to identify a subgroup of head and neck patients that would receive a significantly reduced NTCP if treated with proton therapy rather than X-ray therapy. In their study, patient-specific NTCPs were predicted using proton and photon treatment plans for each patient. Subgroups were defined based on primary tumour location. This allowed groups that would experience a greater NTCP reduction to be identified. The implication is that comparative planning would not be required for all patients in clinics with more limited resources.

2.1.2 Cost-effectiveness of proton therapy

Health care represents a significant portion of government expenditure in many countries. This corresponds to an important investment as the burden of disease is costly for society. Understanding the cost-effectiveness of treatments is vital to ensuring the most appropriate and efficient allocation of limited resources. Proton therapy has the potential to improve quality of life and may even be lifesaving for some patients. However, as discussed in Section 1.2, it has been estimated that proton therapy is 2.5 times more expensive than X-ray therapy. In many instances, this cost may be justified if quality of life can be improved, the cost of treating side effects is reduced, or if the patient is more productive in society as a result of being disease free. Therefore, it is important to consider cost-effectiveness when selecting patients for proton therapy. Cost-effectiveness studies typically involve comparing the costs and outcomes or benefits of two treatments.

Verma *et al.* [38] identified factors that affect the cost-effectiveness of proton therapy. These include patient age, risk of toxicity, and tumour characteristics. Proton therapy is likely to be increasingly cost-effective where tumours are located in close proximity to dose-limiting organs at risk. Alternatively, it may not be cost-effective to treat tumours that have poor prognoses with proton therapy as it is unlikely that patients will live long enough to experience lower toxicity, unless potential side effects will impact significantly on quality of life or mortality risks. Verma *et al.* [39] concluded proton therapy to be cost-effective for paediatric brain tumours, selected breast cancers, locoregionally advanced non-small cell lung cancer and high risk head and neck cancers. However, it was not

demonstrated that prostate cancer or early stage non-small cell lung cancer can be treated with proton therapy cost-effectively in the review.

Lundkvist *et al.* [40–42] investigated the cost-effectiveness of treating various cancers with proton therapy. They employed a Markov model to estimate patient outcomes, which are commonly used in cost-effectiveness studies with medical applications (discussed in more detail in Section 2.2). The findings of this cost-effectiveness research suggest that proton therapy can be cost-effective and cost-saving if appropriate patients are selected to receive the treatment. Childhood medulloblastoma may be treated cost-effectively, particularly as a result of costs associated with hormone replacement therapy and with lost productivity due to IQ reduction. Lundkvist *et al.* [41] recommended that only the cases where there is a higher risk of growth hormone deficiency and IQ loss with X-ray therapy would be cost-effective proton treatments, highlighting a need for individualised patient selection strategies.

Mailhot Vega *et al.* [43] also developed a Markov model to evaluate the cost-effectiveness of treating childhood brain tumours with proton therapy with regard to risk of growth hormone deficiency, and of childhood medulloblastoma specifically [44]. Proton therapy was found to be a cost-effective treatment for medulloblastoma where there were a wide range of potential toxicities including coronary artery disease, congestive heart failure, ototoxicity, gonadotropin deficiency, growth hormone deficiency, hypothyroidism, adrenocorticotrophic hormone deficiency, secondary malignancy, relapse and death. In this study, the risk of growth hormone deficiency was found to be the most influential factor in determining the cost-effectiveness of a proton treatment. It was subsequently concluded [43] that patients who have the greatest potential for sparing of the hypothalamus with proton therapy would be the best candidates to receive the treatment in terms of cost-effectiveness. Proton therapy could also be cost-saving with increased sparing of this tissue.

While paediatric and brain cancers are typically regarded as standard indications for proton therapy, this is not the case for breast cancer. Treating left-sided breast cancer with proton therapy has been found to be cost-effective but only for selected patients with a high risk of developing cardiac disease [40]. The Markov model developed by

Mailhot Vega *et al.* [45] produced similar predictions to Lundkvist *et al.* [40] in that proton therapy was found to be cost-effective for patients who had an elevated risk of cardiac problems. In addition, it was found that patients with these risk factors who would receive a mean heart dose of greater than 5 Gy when treated with photons could be treated with proton therapy cost-effectively, and should therefore be prioritised. The cost-effectiveness of proton therapy may also increase as a wider range of indications are identified [42].

Ramaekers *et al.* [46] analysed the cost-effectiveness of treating head and neck cancers with swallowing sparing photon and proton therapy. Swallowing sparing treatments for head and neck cancer were compared considering risks of dysphagia (swallowing difficulty) and xerostomia (reduced saliva production). If equal survival between the two modalities is assumed, then proton therapy would be cost-effective for selected patients only.

2.1.3 Summary

Decision making in radiation oncology is an important and complex task. Several factors must be considered, including patient quality of life and treatment cost-effectiveness. In the absence of data from Phase III clinical trials of proton therapy, model-based approaches have been proposed to provide evidence as to the most appropriate use of the treatment. The limitation of the Proton Priority System [36] is that only clinical variables are included and there is no consideration of patient-specific dosimetry. The patient selection strategy employed in the Netherlands [28] and the system proposed by Mailhot Vega *et al.* [43] are based on a limited number of NTCP models and provide a quantitative comparison of treatment plans. These models incorporate effects of individual patient dosimetry and clinical variables. However, there is no consideration of the effects of treatment failure or radiation-induced cancers on patient quality of life. Furthermore, cost-effectiveness, which can be important where a treatment has limited availability and is provided with public funding, is not typically included in patient selection systems (with the exception of Mailhot Vega *et al.* [43]). Cost-effectiveness studies have indicated that selected patients could be treated cost-effectively. However, many of these studies have relied on cohort-based estimates of toxicity risks and do not consider the dosimetry

that is unique to the treatment plans of an individual patient. As a result, there remains potential for more detailed modelling approaches.

In this thesis, models are developed that include the effects of treatment failure and second primary cancer induction on patient quality of life, as well as a range of toxicity risks. These effects are assessed using patient-specific dosimetry and, collectively, facilitate a more complete representation of treatment outcomes. Furthermore, treatment cost-effectiveness is incorporated into this work to allow the potential benefit of a treatment to be evaluated from an economic perspective. This is particularly important in the context of healthcare delivery, where resources are often limited. As an additional enhancement, individualised cost-effectiveness predictions are possible, which are based on the patient-specific dosimetry resulting from a given radiation treatment plan.

2.2 Markov models

A Markov model, or Markov chain, is a type of mathematical model used to model randomly-varying systems. A theoretical background of these models is provided by Grimmett and Stirzaker [47]. An important property of Markov models is the *memoryless* Markov property whereby future events depend on current events only, and are independent of past history. The variable of time can be either continuous or discrete. Discrete-time Markov chains are considered here. At each time point, an event may or may not occur, based on a given probability.

2.2.1 Theoretical background

Let $\{X_i\}_{i \in \mathbb{N}}$, where \mathbb{N} is the set of natural numbers, be a sequence of discrete random variables that take values from the finite set S , which is known as the state space of the chain. Each value of S may be thought of as a state of a randomly changing system. If $x_0, x_1, \dots, x_{n-1}, s \in S$, then the Markov property can be written as:

$$\Pr(X_n = s | X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1}) = \Pr(X_n = s | X_{n-1} = x_{n-1}), \quad (2.1)$$

for all $n \geq 1$. The evolution of a chain with time is determined by transition probabilities, $\Pr(X_{n+1} = j | X_n = i)$. These are the probabilities of the system transitioning from being in state i at a given time to state j at the subsequent time point, for $i, j \in S$. A chain is said to be *homogeneous* if:

$$\Pr(X_{n+1} = j | X_n = i) = \Pr(X_1 = j | X_0 = i), \quad (2.2)$$

for all $n \geq 0$, that is, the probability of transitioning in one step between two given states is constant with time. If the transition probabilities are time variable, then the chain is *inhomogeneous*.

All transition probabilities may be represented by a $|S| \times |S|$ matrix, \mathbf{P} , where $|S|$ is the number of Markov states and each (i, j) element of \mathbf{P} is defined as

$$p_{ij} = \Pr(X_{n+1} = j | X_n = i). \quad (2.3)$$

Each row of the transition matrix must sum to 1. The values of \mathbf{P} change at each time point for an inhomogeneous chain.

2.2.2 Markov models in radiotherapy

Markov models have been applied in the discipline of radiation oncology [48, 49]. They may be used to approximate disease progression over a period of time, particularly after a choice of disease management strategy. Decision support systems are often based on Markov models. The Assessment of New Radiation Oncology Technology and Treatment (ANROTAT) Project [50] (undertaken by the Tasman Radiation Oncology Group (TROG)) has developed a framework Markov model to aid the assessment of which new technologies should be funded by Medicare. The Markov model allows the cost-effectiveness of technologies to be evaluated, and hence technologies are selected based on their potential to maximise overall societal benefit.

Sonnenberg and Beck [51] provided a detailed description of the theoretical basis of Markov models applied to medical decision making. At any given time, a patient is

assumed to exist in one of a finite number of discrete states. Each state in the state space describes the health status of the patient. For example, if the patient is disease free, then they are in the Well state. If pneumonia was present, for example, then they would occupy the Pneumonia state.

The time period of interest is a portion of the patient's lifetime. The *cycle length* is the period of time between consecutive discrete time points. In medical applications, the cycle length is typically one year. At the end of each cycle, it is possible for the patient to transition to another health state. Figure 2.2 shows an example of this process. The series of these transitions can be modelled using a Markov chain. Due to the Markov property, the probability of the patient transitioning to another state depends only on their current state and not on the time spent in any previous states. Markov state transition diagrams are used to represent the possible transitions between states. A simple state transition diagram is given in Figure 2.3.

The Dead, or Deceased, state is particularly important as it is the absorbing state. The patient cannot leave this state and this allows the Markov chain to terminate. The *time horizon* is the maximum possible value that the time variable can take. The chain terminates at this point if the absorbing state has not been reached prior. In this work, a time horizon of 100 years is used.

An example of a Markov model utilised in radiation oncology is the decision aid developed by Smith *et al.* [31] (discussed in Section 2.1), which is based on both a Markov model and Bayesian network. Another example of a Markov model was developed by Punglia *et al.* [49] to simulate the progression of ductal carcinoma *in situ* after radiation therapy. Two treatment strategies, excision and radiotherapy and excision alone, were evaluated with the model. The analysis considered local recurrence, breast cancer mortality and mastectomy risks. The model predicted that radiation therapy resulted in a slight improvement of disease-free and overall survival compared with excision alone. However, radiation therapy was also associated with higher probability of mastectomy. This highlights the importance of including patient preferences in decision tools.

Markov models are also often utilised in cost-effectiveness studies (see Section 2.1.2). Lundkvist *et al.* [40] developed such a model to assess the cost-effectiveness treating

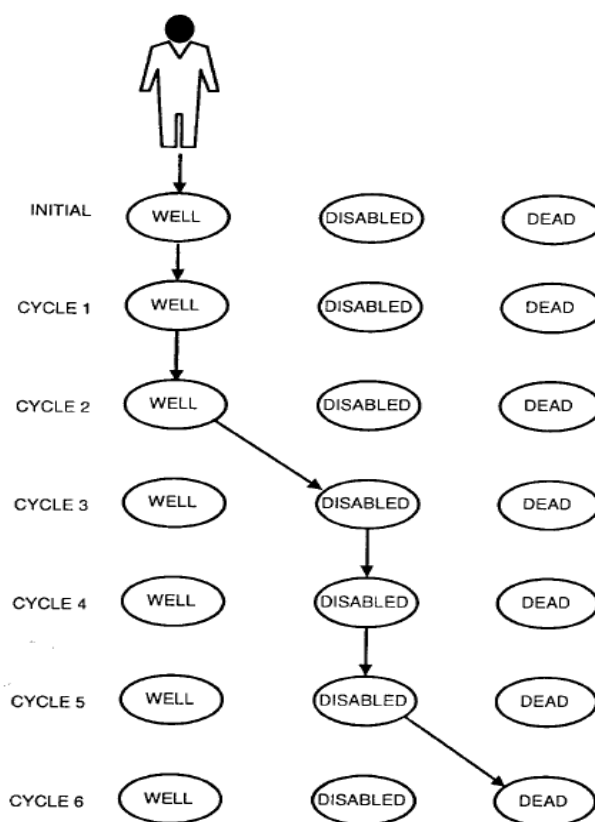


FIGURE 2.2: State transitions made by a person in a Markov chain until the Dead state is reached in cycle 6. Source: [51].

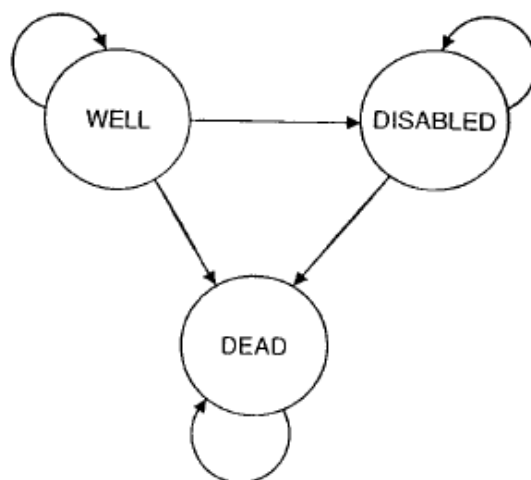


FIGURE 2.3: A simple Markov state transition diagram. There are three states represented by circles. Allowed transitions are represented by arrows. Where an arrow leads from a state to itself again, a patient is allowed to remain in that state in consecutive cycles. Note that recovery from disability is not possible in this example. Source: [51].

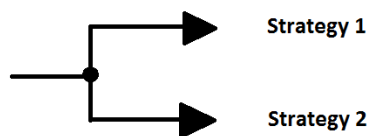


FIGURE 2.4: A simple decision tree diagram. The decision node is found at the centre. The outcome nodes are found at the end of the branches representing each strategy.

left-sided breast cancer with proton therapy. Ramaekers *et al.* [46] (introduced in 2.1.2) also utilised Markov modelling to investigate proton therapy cost-effectiveness for the treatment of head and neck cancer.

2.2.2.1 Comparison with other prognosis modelling approaches

Decision tree models consist of three components as illustrated by Figure 2.4 [52]: (i) the decision node which is the point in time when a choice is made between two treatment strategies; (ii) the decision strategies which correspond to the branches of the tree; and (iii) the outcome nodes which are the outcomes of the strategies. For the simple tree in Figure 2.4, it is not possible to specify when events occur, or for events to occur more than once. While additional trees could be added to outcome nodes to address this, the result would be an unacceptably high number of branches. Markov models offer a much more convenient means of tracking the timing of events, which is particularly important for estimating prognoses [51].

Discrete-event simulations (DES) differ from Markov models in that they are event oriented rather than state oriented [53]. DES can have the advantage of being able to more simply model patient histories and competing risks [54]. The time between events is not necessarily constant in DES. Events may be instantaneous and are not mutually exclusive. The memoryless property of Markov models means that it is not possible to easily include patient history without significantly increasing the number of Markov states. However, the disadvantage of DES is that large datasets are often required which may not always be available.

2.2.3 Quality of life concept and utilities

A Markov process can be used to model the life expectancy of a patient by counting the number of cycles that pass before the Deceased state is reached. However, the quality of a patient's life may also be of interest, and this is affected by cancer or injuries arising from treatment complications. The quality-adjusted life expectancy (QALE) is an adjustment of the raw life expectancy to account for these factors. To calculate the QALE, each state is assigned a quality of life (QoL) utility (also known as a QoL weight). The QALE is given by (2.4), where t_i is the total number of cycles the patient spends in the i^{th} state, $u_i \in [0, 1]$ is the QoL value for the i^{th} state, and n is the total number of possible states.

$$\text{QALE} = \sum_{i=1}^n t_i u_i. \quad (2.4)$$

The QALE may also be thought of as the number of quality-adjusted life years (QALYs) lived until death. The advantage of the model that has been developed in this work is that the probabilities of tumour control and of developing complications are incorporated into a single metric, the QALE, which is used to compare treatment plans. By incorporating both these factors into the analysis, a more informed comparison can be made between two treatment plans.

The value of the utility depends only on the associated state and not history of health state occupation. The Well state has $u = 1$, Dead has $u = 0$, and the other states have a value in $(0, 1)$ [55]. A variety of methods may be used to assign utilities to the remaining states. For example, Kharroubi *et al.* [56] used Bayesian Markov chain Monte Carlo methods and clinical data to estimate utilities that describe a patient's quality of life. The Bayesian models also provide uncertainty distributions associated with these values. The utilities estimated were intended for use in cost-effectiveness analyses [56]. Sonnenberg and Beck [51] also noted that it is possible to enhance Markov models using time-dependent utility values.

2.2.4 Summary

Markov models provide a framework for modelling disease progression and patient quality of life. While they have limitations, these models have advantages compared with other methods that are applied in prognosis modelling. For example, they may prove useful where data availability is limited. Markov approaches have been utilised for investigating clinical outcomes of various radiation treatments, including proton therapy. The results of these investigations highlight the need to select appropriate patients to receive proton therapy. Therefore, Markov models may also prove useful when implemented as part of proton therapy patient selection strategies.

2.3 Radiobiological models

Ionising radiation has the potential to cause damage to biological tissue, specifically through damaging DNA. A variety of mathematical models exist that relate the radiation dose to a clinical endpoint. Several of these radiobiological models were incorporated into this work as dose-dependent transition probabilities in the Markov model. The endpoints considered included tumour control, normal tissue complications and second malignancies (radiation-induced). The radiobiological models employed to calculate the probabilities of these endpoints occurring are introduced and discussed in this section.

Other studies have utilised alternative methods to determine Markov transition probabilities. The transition probabilities in the Markov model utilised by Lundkvist *et al.* [40] were estimated based on toxicity and mortality rates available in the literature, rather than based on patient-specific doses. Smith *et al.* [31] derived transition probabilities from a Bayesian network.

Radiobiological models have also been used to compare treatment plans directly, without using them as input to a Markov model. The proton therapy patient selection approach proposed by Langendijk *et al.* [28] involved calculating the NTCP using the dose to a given organ resulting from a treatment plan, along with other clinical variables. The NTCP resulting from a photon plan is compared with the NTCP expected with a proton

plan. Hence, the NTCP is used to quantify the clinical outcome. The disadvantage of this approach is that this quantification does not consider other factors such as the induction of second cancers. The advantage of Markov models is that several factors may be combined to give a single metric that can indicate the quality of a plan.

2.3.1 TCP

In this work, the tumour control probability (TCP) is used in the model to determine whether a radiotherapy treatment is successful. As a result, it is an important parameter as it determines the state that a patient begins the Markov simulation in (either Well for a successful treatment or Cancer for an unsuccessful treatment). The TCP is based on the linear quadratic (LQ) model [57],

$$S(D) = \exp(-\alpha D - \beta D^2), \quad (2.5)$$

which gives the fractional number of surviving cells after being irradiated by a certain dose, D , where α and β are the linear and quadratic coefficients of the LQ model, respectively. In general, TCP models assume that the number of surviving tumour cells follows a Poisson distribution and that a single surviving clonogen (cancerous cell) is required for tumour regrowth. Hence tumour control corresponds to all cells being killed. The total TCP is given as a product over the tumour's volume elements (or voxels) [58, 59]:

$$\text{TCP} = \prod_{i=1}^{\ell} P(D_i)^{v_i}, \quad (2.6)$$

$$P(D_i) = \exp\left(-\exp\left(e\gamma - \alpha D_i - \beta \frac{D_i^2}{n}\right)\right), \quad (2.7)$$

where there are a total of ℓ voxels, with each having a fractional volume (of the total tumour) v_i that receives dose D_i as part of a treatment delivered in n fractions. The parameter $\gamma = D \frac{d\text{TCP}}{dD}$ is the normalised dose-response gradient evaluated at $D = D_{50}$, the treatment dose D at which 50% of tumours are controlled.

2.3.2 NTCP modelling

The normal tissue complication probability (NTCP) represents the probability of developing a radiation-induced injury as a result of treatment. NTCP models are also based on the dose received by volume fractions of a particular organ. Depending on the properties of a given organ, it may respond as a serial or parallel organ. However, there is a spectrum of organ behaviour, with purely serial and purely parallel organs representing opposite ends of the spectrum. The response of a purely parallel organ, the lung for example, has a greater dependency on the irradiated volume. For example, if a larger volume is irradiated, then there will be a greater impact on organ function. In contrast, response of a purely serial organ, the spinal cord for example, has a greater dependency on the magnitude of the dose rather than the irradiated volume. The functioning of these organs can be impaired by irradiation of a small volume. Therefore, NTCP models should incorporate the effect of organ volume dependency.

2.3.2.1 LKB model

The Lyman-Kutcher-Burman (LKB) NTCP formalism is given by ([60], [61])

$$\text{NTCP} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t \exp\left(\frac{-x^2}{2}\right) dx, \quad (2.8)$$

$$\text{with } t = \frac{D_{eff} - TD_{50}}{mTD_{50}} \quad \text{and} \quad D_{eff} = \left(\sum_{i=1}^{\ell} v_i D_i^{\frac{1}{n}} \right)^n,$$

where TD_{50} is the uniform dose given to the entire organ that results in 50% complication risk, m is an organ specific parameter that is related to $\frac{d\text{NTCP}}{dD}$, n is a parameter that characterises the volume dependence of the organ's response to radiation, D_{eff} is the effective dose, and ℓ is the number of voxels. For a purely parallel organ, $n = 1$ and for a purely serial organ, $n = 0$.

The parameters in the model are complication-specific. Ideally, these are clinically founded. An advantage of the LKB model is that there are many published parameters for a wide range of endpoints [62–64]. The Quantitative Analysis of Normal Tissue

Effects in the Clinic (QUANTEC) data is an important resource for these parameters [65].

However, a limitation of the LKB model is that patient-specific clinical variables such as disease history and concurrent chemotherapy are not included. These variables can have a significant influence on complication rates.

2.3.2.2 Other NTCP models

Other methods of predicting NTCPs exist. The relative seriality model, for example, is given by:

$$\text{NTCP} = \left(1 - \prod_{i=1}^{\ell} [1 - P(D_i)^s]^{v_i} \right)^{1/s}, \quad (2.9)$$

where $s = 1$ for a purely serial organ, $s = 0$ for a purely parallel organ, D_i , v_i and ℓ have the same meaning as for the LKB model, and $P(D_i)$ is given by (2.7). The relative seriality model is very similar in principle to the LKB model; however, the definition of the parameters varies slightly.

NTCP models have been developed that depend on both dose and clinical variables (such as tumour stage, patient age and sex, concurrent chemotherapy). Several such models have been developed in the Netherlands for several endpoints including xerostomia [66, 67], swallowing dysfunction [68] and tube feeding dependency [69]. These have been developed through the method of logistic regression to identify which variables contributed to the value of the NTCP. Many of these models have been implemented as part of the proton therapy patient selection system that has been adopted in the Netherlands [28].

External validation is important to ensure the reliability of these models. This process involves comparing the predictions of the model with observed outcomes of a treatment cohort that is independent of the cohort utilised to develop the model. Several NTCP models exist have been subjected to this rigorous testing process [70–72].

2.3.3 Second cancer risk

The second primary cancer induction probability (SPCIP) is the probability of developing a cancer as a result of the treatment radiation. This is distinct from recurrence or metastasis of the primary treated cancer, which is referred to as secondary cancer in this work. This is a particularly important consideration for paediatric patients as radiation-induced cancers can take several years, even decades, to develop [73]. Therefore, it is more likely that older patients would be less concerned by the threat posed by these cancers, while younger patients have a longer remaining life-time over which to develop second cancers. In addition, younger patients have a greater proportion of proliferating cells, which are more likely to be in a radiosensitive (more susceptible to radiation damage) stage of the cell cycle [74]. Younger patients are also more likely to be affected by rare tumours that are typically close to critical structures in the central nervous system [75]. Therefore, proton therapy is often more appropriate for younger patients than photon treatments. As a consequence, it is important to consider the effect of second cancers on patient quality of life when selecting patients for proton therapy.

Comparative planning studies have indicated that proton therapy is associated with a significantly reduced risk of second cancers. Examples include patients with liver metastases receiving proton beam radiosurgery or photon beam radiosurgery [76], and medulloblastoma patients receiving spinal irradiations with IMPT (lifetime risk of 4%) or IMRT (lifetime risk of 30%) [77]. The latter could be explained by the greater risk of second malignancies associated with IMRT compared with conventional X-ray treatments [78], [79]. In contrast, observations of second malignancy incidences after proton therapy indicate that the treatment is not associated with a significantly increased risk of secondary malignancies compared with photon therapy [80]. However, these observations were over a relatively short follow-up period (median 6.7 years) compared with the time periods over which these cancers can often develop [73].

Malignant transformation of cells is a stochastic process and usually occurs in the lower dose region whereas cell death (leading to injury) occurs in the higher dose region. Therefore, radiation that has been scattered during the treatment delivery is more likely to result in a second cancer due to the lower dose it delivers [73]. While proton therapy

reduces the integral dose to a patient, thus reducing the probability of cancer induction [81], the out-of-field doses are not negligible. This further highlights the need to consider the risk of second cancers when comparing proton and photon therapies [82].

Timlin *et al.* [83] developed a model to predict malignant induction probabilities for a given dose to an individual patient, while considering varying responses of different tissue types. They proposed that their model could be used for comparing radiation treatment plans, but noted that validation is required to allow meaningful comparisons.

In this work, the SPICP is based on the model of Schneider *et al.* [84]. Here, the volumes that receive a given dose are used to calculate the excess absolute risk (EAR) of cancer induction for each year after exposure. Note that this is a cumulative probability and differential probabilities were defined for this work as the difference in EAR between adjacent years. The EAR for a particular organ at a particular time after treatment due to radiation exposure is given by (2.10) [84]

$$EAR^{org}(age) = \frac{1}{V_T} \sum_{i=1}^{\ell} v_i(D_i) \beta_{EAR} RED(D_i) \mu(ageX, age). \quad (2.10)$$

Here, $ageX$ is the age of the patient at the time of treatment (the time of exposure to radiation), age is the age of the patient after treatment at the year of interest, V_T is the total volume of the organ, β_{EAR} is the initial slope of the dose-response curve, ℓ is the total number of voxels, and

$$\mu(ageX, age) = \exp \left[\gamma_e (ageX - 30) + \gamma_a \ln \left(\frac{age}{30} \right) \right], \quad (2.11)$$

with γ_e and γ_a being the age modifying parameters. The EAR model parameters assume an age at exposure of 30 years and the age modifying parameters account for this. The function $RED(D)$ is the risk equivalent dose (RED) mechanistic model which accounts for the effects of cell killing and fractionation,

$$RED(D) = \frac{e^{-\alpha'D}}{\alpha'R} \left(1 - 2R + R^2 e^{\alpha'D} - (1 - R)^2 e^{\frac{\alpha'R}{1-R} D} \right), \quad (2.12)$$

where R is the repopulation/repair parameter, α' is given by

$$\alpha' = \alpha + \beta \frac{D}{D_T} d_T, \quad (2.13)$$

and D_T and d_T represent the prescribed dose to the target volume and the corresponding dose per fraction, respectively.

The advantage of this model is that it combines data from the low dose region (A-bomb survivors) with data from the higher dose region (Hodgkin's Disease patients) to produce a more reliable prediction in dose regions relevant to radiotherapy patients. In addition, the effect of radiation treatment fractionation (where the treatment dose is delivered in small, regular fractions) is incorporated. The limitation is the lack of clinical validation, which can be difficult due to the long follow-up period required.

2.3.4 Summary

Radiobiological models have the power to quantify the complex relationship between radiation dose and biological effect. Several models have been developed to quantify a variety of endpoints. However, it is not advisable to implement these models clinically without proper validation, otherwise a large degree of uncertainty exists regarding the accuracy of their predictions. Radiation-induced malignancies are a particularly important consideration when comparing photon and proton treatments at the level of individual patients. Validation of these models can be practically challenging due to the long follow-up periods required. These uncertainties and limitations will be discussed in the publications that form the body of this thesis.

Chapter 3

Monte Carlo Evaluation

The publication P1 forms the basis of this chapter.

Austin, A.M., Douglass, M.J.J., Nguyen, G.T. & Penfold, SN. A radiobiological Markov simulation tool for aiding decision making in proton therapy referral. *Physica Medica*. 2017; 44:72–82.

3.1 Introduction and motivation

In this chapter, the development of the first patient selection model is described. The distinguishing feature of this model is that it is based on a Monte Carlo simulation. This preliminary model served as the basis for future patient selection models. However, this model is complete and is able to predict the clinical outcome of an individual patient from a given treatment plan.

The aims of this publication were to introduce the toolkit and to demonstrate the output with an example patient, thus quantitatively determining whether the patient in question would benefit from proton therapy.

Technical details relating to the development of this model are also presented in this chapter, along with the results of the model verification. This verification was carried out to ensure that the model was behaving as expected.

3.2 Statement of contribution

3.2.1 Conception

The idea to use a Markov model as a tool for patient selection for proton therapy was first conceptualised by Scott Penfold. All authors contributed to the development of ideas and methods.

3.2.2 Realisation

The writing of the code and analysis was performed by Annabelle Austin, with advice provided by Scott Penfold, Michael Douglass and Giang Nguyen.

3.2.3 Documentation

This paper was written by Annabelle Austin. Editing was performed by all authors.

Statement of Authorship

Title of Paper	A radiobiological Markov simulation tool for aiding decision making in proton therapy referral
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Austin, A.M., Douglass, M.J.J., Nguyen, G.T. & Penfold, SN. A radiobiological Markov simulation tool for aiding decision making in proton therapy referral. <i>Physica Medica</i> . 2017; 44:72–82.

Principal Author

Name of Principal Author (Candidate)	Annabelle Austin		
Contribution to the Paper	Developed code for the tool, performed simulations and analysis, wrote manuscript and acted as corresponding author.		
Overall percentage (%)	85%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	27/05/19

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Scott Penfold		
Contribution to the Paper	Provided supervision, advice, and assisted in manuscript evaluation.		
Signature		Date	22/05/19

Name of Co-Author	Michael Douglass		
Contribution to the Paper	Provided supervision, advice, and assisted in manuscript evaluation.		
Signature		Date	22/05/19

Name of Co-Author	Giang Nguyen		
Contribution to the Paper	Provided supervision, advice and assisted in manuscript evaluation.		
Signature		Date	24/05/2019



Original paper

A radiobiological Markov simulation tool for aiding decision making in proton therapy referral



Annabelle M. Austin^{a,*}, Michael J.J. Douglass^{a,b}, Giang T. Nguyen^c, Scott N. Penfold^{a,b}

^a Department of Physics, University of Adelaide, Adelaide, SA 5005, Australia

^b Department of Medical Physics, Royal Adelaide Hospital, Adelaide, SA 5000, Australia

^c School of Mathematical Sciences, University of Adelaide, Adelaide, SA 5005, Australia

ARTICLE INFO

Keywords:

Proton therapy
Markov model
Decision aid
Patient selection

ABSTRACT

Purpose: Proton therapy can be a highly effective strategy for the treatment of tumours. However, compared with X-ray therapy it is more expensive and has limited availability. In addition, it is not always clear whether it will benefit an individual patient more than a course of traditional X-ray therapy. Basing a treatment decision on outcomes of clinical trials can be difficult due to a shortage of data. Predictive modelling studies are becoming an attractive alternative to supplement clinical decisions. The aim of the current work is to present a Markov framework that compares clinical outcomes for proton and X-ray therapy.

Methods: A Markov model has been developed which estimates the radiobiological effect of a given treatment plan. This radiobiological effect is estimated using the tumour control probability (TCP), normal tissue complication probability (NTCP) and second primary cancer induction probability (SPCIP). These metrics are used as transition probabilities in the Markov chain. The clinical outcome is quantified by the quality adjusted life expectancy. To demonstrate functionality, the model was applied to a 6-year-old patient presenting with skull base chordoma.

Results: The model was successfully developed to compare clinical outcomes for proton and X-ray treatment plans. For the example patient considered, it was predicted that proton therapy would offer a significant advantage compared with volumetric modulated arc therapy in terms of survival and mitigating injuries.

Conclusions: The functionality of the model was demonstrated using the example patient. The proposed Markov method may be a useful tool for deciding on a treatment strategy for individual patients.

1. Introduction

The use of intensity modulated proton therapy (IMPT) for the treatment of cancer has become increasingly common in recent years. The primary advantage of IMPT over intensity modulated radiation therapy with X-rays (IMRT) lies in reduction of integral dose deposited in the patient while delivering an equivalent dose to the tumour volume [1,2]. The disadvantage of proton therapy is that it is a more expensive form of treatment with limited availability.

The issue of limited availability suggests that proton therapy should be prescribed for those patients who will benefit most when compared to treatment with conventional X-ray therapy. However, it is often difficult to base a treatment decision for a given patient on the results of randomised Phase III clinical trial data comparing novel and standard treatments. One of the main issues is the long follow-up times required for these clinical trials in an environment of rapidly evolving

radiotherapy technology, with results potentially becoming obsolete shortly after they are gathered.

The concept of *in silico* clinical trials for proton therapy was proposed by Langendijk et al. [3] to address this issue. Langendijk et al. [3] suggest the use of normal tissue complication probabilities (NTCPs) as a discriminator for when a patient should receive proton therapy. Using validated, evidence-based parameters for the calculation of a given NTCP, a comparison of values for proton therapy and X-ray therapy is made. If the NTCP is reduced by a certain threshold value in the proton plan relative to the X-ray plan, the patient is eligible to receive proton therapy. A key assumption of this model is that the tumour control probability (TCP) is equal for both treatments and second primary cancer induction probabilities (SPCIPs) are neglected. Considering the difference in integral dose one can expect when comparing an X-ray treatment plan with a proton treatment plan, the latter assumption warrants further investigation. This is particularly true for paediatric

* Corresponding author.

E-mail addresses: annabelle.austin@adelaide.edu.au (A.M. Austin), michael.douglass@adelaide.edu.au (M.J.J. Douglass), giang.nguyen@adelaide.edu.au (G.T. Nguyen), Scott.Penfold@sa.gov.au (S.N. Penfold).

<https://doi.org/10.1016/j.ejmp.2017.11.013>

Received 1 September 2017; Received in revised form 7 November 2017; Accepted 15 November 2017

Available online 23 November 2017

1120-1797/ Crown Copyright © 2017 Published by Elsevier Ltd on behalf of Associazione Italiana di Fisica Medica. All rights reserved.

cancers where the patient may potentially live for many years after treatment. A review of the potential use of radiation-induced cancer risk predictive models was presented by Stokkevaåg et al. [4]. Another example of work comparing X-ray and proton therapy (specifically IMRT and spot-scanning proton therapy), based on NTCP modelling only was presented by Yoshimura et al. [5].

In the current work we propose an *in silico* clinical trial model which accounts for NTCP as well as TCP and SPCIP. To achieve this, a Markov simulation framework was developed which aims to combine the dosimetric data from all contoured structures to provide an estimate of the quality adjusted life expectancy (QALE) resulting from a given treatment plan. A Markov model is a type of stochastic model that is commonly applied in medical decision making [6–8]. The International Society for Pharmacoeconomics and Outcomes Research Modelling Good Research Practices Task Force has prepared a report outlining best practices for developing and implementing Markov models in medical applications [9]. The core quantitative output of the Markov model presented in the current work is the QALE, which is an adjustment of the raw life expectancy to account for the effect of poor health on the quality of life of the patient. Therefore, estimating the QALE with a Markov model can allow a quantitative comparison of treatment plans at the level of clinical outcome. An example of a Markov model applied to proton therapy was presented by Ramaekers et al. [10], who developed a tool for comparing proton and photon treatments for head and neck cancers based on NTCP models. Xerostomia and dysphagia were the only two toxicities considered which were assumed to be irreversible after the first six months.

The novelty of the Markov framework presented in the current work lies in the inclusion of a wide range of NTCPs and also the effects of TCP and SPCIP on a patient's quality of life. The induction of second cancers is an important consideration for the younger patients often treated with proton therapy. Separate NTCPs and SPCIPs were considered for various organs at risk (OARs) to allow for a more realistic model, by distinguishing between the effects of different treatment complications. Several potential toxicities were included in addition to xerostomia and dysphagia. As a further enhancement of the work of Ramaekers et al. [10], all transition probabilities are time-variable in the proposed model.

As a demonstrative test case, the model was used to estimate the outcomes for a paediatric base of skull chordoma (BOSCh) patient after receiving an IMPT and a volumetric modulated arc therapy (VMAT) treatment. BOSCh is a very rare form of bone tumour that is difficult to treat with radiotherapy due to the proximity of the spinal chord. In addition, this type of tumour typically affects younger patients who have a both a heightened sensitivity to radiation-induced cancer and longer remaining lifetime over which to potentially develop one. Proton therapy is seen as an advantageous treatment for this disease as it delivers a much smaller integral dose to the body compared with IMRT [11]. The intention of this test was to demonstrate the capabilities of the model rather than compare the efficacy of IMRT and IMPT for the treatment of BOSCh.

In summary, the aim of this work was to develop and present a Markov framework that included TCP and SPCIP in addition to NTCP models to estimate the clinical outcome of a proton or X-ray treatment plan on an individual patient basis. The technical details of the model are presented in Section 2 with a description of the Markov states, transition probabilities and model verification. The clinical example and results are described in detail in Section 3 and a discussion and conclusion is given in Sections 4 and 5.

2. The Markov model

With a discrete-time inhomogeneous (time-dependent) Markov chain, the response of a patient to a particular treatment is modelled by approximating the remainder of the patient's life as a series of transitions between a finite number of discrete states. The patient can occupy

only a single state at a given time. These states describe the health status of a patient and include Well, Deceased and the Diseased group of states (Section 2.1). The course of radiation therapy (either proton or photon) is the strategy that is evaluated by the model and hence represents a one-time intervention at the beginning of the Markov chain. There are no subsequent interventions (additional treatments for example) in this model.

In a Markov chain, the time period of interest (the patient's lifetime in this case) is divided into equal increments, or cycles. The cycle length for this Markov chain was chosen to be one year as this provided computational efficiency while allowing the Markov chain to have a large number of cycles. This is particularly important in the context of radiotherapy where complications can arise several years after treatment. The model assumes that the patient remains in a particular state for the duration of a cycle. At the end of each cycle, it is possible for the patient to transition to another state. An important property of a Markov process is memorylessness, which in this context implies that the probability of the patient transitioning to another state depends only on their current state and not on any previous states. The transition probabilities can vary with time and may be different for distinct transitions between state pairs. The patient response is quantified by the QALE, which is the primary metric used to evaluate and compare treatment plans. Kaplan-Meier plots also provide a useful visual summary of the clinical outcome. In order to obtain statistical results, the Markov process is simulated many times with each simulation representing a member of a hypothetical patient cohort. All tables referred to in this section can be found in the Appendix.

2.1. Markov states

Fig. 1 shows the Markov states for the simple case where there is only one OAR being considered. In general however, the model can consider any number of OARs. Each state is represented by a node in Fig. 1. In medical applications of Markov models, the Well state corresponds to perfect health and the Deceased state represents patient death. The other states represent various cases of poor health. In this work, the Diseased group contains states that represent varying numbers and forms of diseases arising as a result of treatment. These states include the cases in which a patient:

- still has their initial primary cancer due to unsuccessful treatment (Diseased (primary cancer)), represented by the Cancer node
- develops one or more normal tissue complications as a result of treatment (Diseased (injury)), represented by the Inj node
- develops one or more second primary cancers (SPCs) as a result of treatment (Diseased (second cancer)), represented by the SPC node

There are also other states representing every possible combination

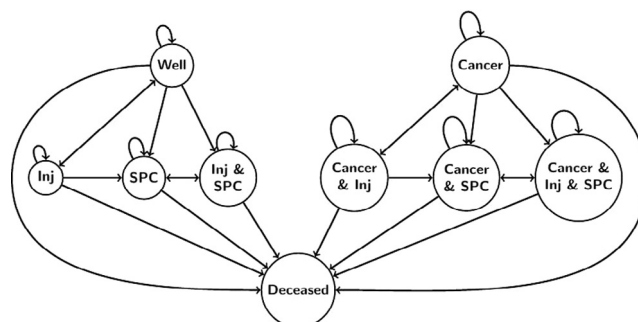


Fig. 1. The Markov state transition diagram showing the allowed transitions between states. For simplicity, this describes the case where there is only one injury and one SPC being considered in the model. “Well” represents perfect health. “Cancer” represents the situation where the patient still has the initial primary cancer and “Inj” represents an injury state.

of the above:

- Diseased (primary cancer and injury), Cancer & Inj node
- Diseased (primary and second cancer), Cancer & SPC node
- Diseased (second cancer and injury), Inj & SPC node
- Diseased (primary and second cancer and injury), Cancer & Inj & SPC node

The number of Markov states in the Diseased group is a variable which depends on the number of second primary cancers (SPCs) and injuries that are being considered in the model. A state exists for every possible combination of SPCs and injuries that could affect a patient at a given time. In addition, there is also a state for every possible combination of SPCs and injuries for the case where the primary cancer remains as a result of unsuccessful treatment. Fig. 1 shows the allowed transitions between states for the case where there is only one injury and one SPC being considered. Depending on the value of the TCP, the patient can begin the Markov process in either the Well state or the Diseased (primary cancer) state (see Section 2.4.1 for an explanation). Therefore, once the Markov process has begun, it is not possible to transition between these two states, as depicted in Fig. 1. Section 2.4 describes how the probabilities of transitioning between these states are derived.

2.2. Quality adjusted life expectancy

The Deceased state is particularly important as it is the absorbing state of the Markov model. The patient cannot leave this state, which allows the Markov process to terminate. The process will also terminate if the patient reaches 100 years of age. This age represents the time horizon of the Markov model. It was assigned a value of 100 years due to the high probability of a patient deceasing before this age. The model can estimate the life expectancy of a patient by counting the number of cycles that pass before the Deceased state is reached, which is equivalent to the number of years until death. However, the quality of a patient's life is also of interest and this is affected by cancer or injuries arising from complications from treatment. The quality adjusted life expectancy (QALE) is an adjustment of the raw life expectancy to account for these factors. The QALE is the most indicative parameter of the quality of a patient plan. To calculate the QALE, each state was assigned a quality of life (QoL) utility. These values represent the quality of life associated with the state relative to perfect health. By default, the Well state is assigned the maximum QoL utility of 1 and the Deceased state is assigned the minimum QoL utility of 0. All other states were assigned a value within this range (Appendix B). The QALE is given by

$$\text{QALE} = \sum_{i=1}^n t_i u_i, \quad (1)$$

where t_i is the total number of cycles the patient spends in state i , u_i is the QoL utility for state i and n is the total number possible states. As the utilities each have a value on the interval [0,1], the QALE cannot be greater than the raw life expectancy. The advantage of this model is that the probabilities of tumour control and of developing complications are incorporated into a single metric, the QALE, which is used to evaluate treatment plans. Balancing these two factors leads to a more effective comparison.

2.3. The Monte Carlo method

In order to accurately model the likely response of a patient to a course of treatment, it is necessary to employ statistical methods. The Monte Carlo approach to Markov modelling makes use of random numbers and repeated sampling from probability distributions to obtain numerical results.

Initially, a single hypothetical patient, also referred to as a Sim, is considered in the model with an initial age equal to the patient's age at the time of treatment. At the beginning of each cycle, a pseudo-random number η is generated and compared to the transition probabilities between different possible states to determine whether a transition will be made. This process involves considering all allowed transitions from the current state. For example, it is not possible to transition from a Diseased (cancer) state to the Well state. All relevant transition probabilities are normalised to yield relative probabilities and then converted to cumulative probabilities, such that the interval [0,1] is divided into n regions representing the relative magnitude of each probability, $[R_1, R_2, R_3, \dots, R_n]$, where n is the number of allowed transitions and $R_n = 1$. If the normalised probability of transitioning to state k is $P(k)$ for example, then $[R_k, R_{k+1}]$ represents the region within the interval [0,1] corresponding to that transition, where $P(k) = R_{k+1} - R_k$. If $R_k < \eta \leq R_{k+1}$ then the Sim moves to state k , where the value of η is constrained by $0 < \eta \leq 1$. The age and state variables are updated and another cycle occurs. Once this Sim reaches the absorbing Deceased state, or if the Sim reaches the age of 100, the process terminates and the QALE (along with other metrics) is stored. This process is repeated for a large hypothetical cohort of these Sims, with each Sim being an identical copy of the individual patient under consideration. The relative uncertainty is proportional to $1/\sqrt{N}$ where N is the cohort size. A large cohort is therefore required to maximise the precision of the results.

2.4. State transition probabilities

The probabilities of a Sim transitioning between particular states were derived from radiobiological models. These models all require the dose-volume data which is the primary input for the Markov model as a whole. The choice of radiobiological model that is applied depends on the pair of states between which a transition is being made.

2.4.1. Tumour control probability

The tumour control probability (TCP) is a single probability used in the model to determine whether the treatment successfully controlled the tumour. It is an important parameter as it determines the state in which a Sim begins the Markov simulation, either Well for a successful treatment or Diseased (primary cancer only) otherwise. If the calculated TCP is less than a pseudo-random number, then the Sim begins the Markov process in the Well state. If it is greater than or equal to the pseudo-random number, then the Sim begins the process in the Diseased (primary cancer only) state. Once the simulation has begun, it is not possible for the Sim to transition between these two states.

The TCP is based on the linear quadratic (LQ) model which gives the fractional number of surviving cells after being irradiated by a certain dose [12]. In general, TCP models assume that the number of surviving tumour cells follows a Poisson distribution, and that a single surviving clonogen (cancerous cell) is required for tumour regrowth. Hence, tumour control corresponds to all clonogenic cells being killed. The total TCP is given as a product over the tumour's volume elements (or voxels):

$$\text{TCP} = \prod_{i=1}^M P(D_i)^{v_i}, \quad (2)$$

$$\text{with } P(D_i) = \exp\left(-\exp\left(e\gamma - \alpha D_i - \beta \frac{D_i^2}{n_{frac}}\right)\right), \quad (3)$$

where there are a total of M voxels, each having a fractional volume v_i (of the total tumour) that receives dose D_i as part of a treatment delivered in n fractions (n_{frac}), and α and β are the linear and quadratic coefficients of the LQ model, respectively. The parameter γ is the normalised dose-response gradient evaluated at $D = D_{50}$, the dose at which 50% of tumours are controlled.

The particular values used for these parameters in the example patient calculation are listed in Table 4 (Appendix A), along with an explanation of their derivation. The sources of these parameters assumed 2 Gy fractions, and hence the input dose was converted to the equivalent dose in 2 Gy fractions (EQD2), with

$$EQD2_i = D_i \frac{1 + \frac{D_i / n_{frac}}{\alpha / \beta}}{1 + \frac{2}{\alpha / \beta}}, \quad (4)$$

before being used as D_i values to calculate the TCP with (2).

2.4.2. Normal tissue complication probabilities

The normal tissue complication probability (NTCP) represents the probability of developing an injury as a result of treatment. In this simulation tool, there is a separate NTCP for each injury being considered. For example, when treating BOSCh, necrosis of the brainstem, cataracts of the eye lens, and spinal cord myelitis are some of the possible complications that can arise after the commencement of radiotherapy. Distinguishing between different injuries is important as they may have varying effects on a patient's quality of life. The NTCP for a particular injury corresponds to a transition to the Diseased state representing that injury.

The Lyman-Kutcher-Burman NTCP formalism is given by (5) and (6) [12–14],

$$NTCP = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t \exp\left(-\frac{x^2}{2}\right) dx, \quad (5)$$

$$\text{with } t = \frac{D_{eff} - TD_{50}}{mTD_{50}} \text{ and } D_{eff} = \left(\sum_i v_i D_i^{\frac{1}{n}} \right)^n, \quad (6)$$

where TD_{50} is the uniform dose given to the entire organ that results in 50% complication risk, m is an organ-specific parameter that represents the gradient of the dose-response curve (analogous to γ), and n is a parameter that characterises the volume dependence of the organ's response to radiation. Using the same method outlined in Section 2.4.1, the D_i values were obtained by converting the raw dose values to equivalent doses in 2 Gy fractions. The NTCP calculation requires clinically founded organ-specific parameters (listed in Table 5).

Unlike the TCP which is a single number, the NTCP is a time-dependent probability and hence is given by an array of probabilities (one value for each year after treatment). The all-time NTCP is given by (5). The NTCP for each year after treatment is calculated based on a normal distribution. This process involves several steps:

1. A time interval between two years x_1 and x_2 is selected.
2. The probability density function of the normal distribution is integrated using the trapezoidal rule,

$$\int_{x_1}^{x_2} f(x) dx \approx \frac{1}{2} s (f(x_1) + f(x_2)) + s \sum_{k=1}^{N-1} f(x_1 + ks), \quad (7)$$

to give the probability of developing an injury in a particular year between years x_1 and x_2 , where $s = (x_2 - x_1)/N$ is the step size, N is the number of integration steps and the normal density function with mean μ and standard deviation σ is given by

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (8)$$

The parameters μ and σ are estimates of the mean and standard deviation of time taken after treatment for the particular injury to develop (listed in Table 6). An important source of these parameters is the Quantitative Analysis of Normal Tissue Effects in the Clinic (QUANTEC) data [15]. This is a large and comprehensive review of available dose-response data for a variety of normal tissue endpoints. An alternative is the publication ALERT – Adverse Late Effects of Cancer Treatment [16]. Similarly to QUANTEC, it provides a

review of organ specific complications that can arise from treatment, as well as organ specific dose-volume relationships.

3. The probability of developing the injury in a particular year is then scaled so that the total area beneath the normal distribution over the remaining possible life time of the patient (the difference between 100 and the initial age of the patient in years) is equal to the all-time NTCP calculated from (5) for that injury. This calculation is repeated for each injury and the result is an NTCP for each possible injury for each year after treatment. The total all-time NTCPs for each injury are listed in Table 11.

2.4.3. Second primary cancer induction probabilities

The second primary cancer induction probability (SPCIP) represents the probability of developing a second primary cancer as a result of treatment. Similarly to the NTCP, there is a separate SPCIP for each tissue where an SPC could form, which is used in the Markov model as the probability of transitioning to a Diseased state representing that particular SPC. In addition, the SPCIP is a time-dependent probability and has a different value for each year after treatment.

The excess absolute risk (EAR) of developing a cancer in a particular organ at a particular time after treatment due to radiation exposure is given by (9) and was taken as an estimate of the SPCIP [17,18]

$$EAR^{org}(age) = \frac{1}{V_T} \sum_i v_i(D_i) \beta_{EAR} RED(D_i) \mu(ageX, age). \quad (9)$$

Here, $ageX$ is the age of the patient at the time of treatment (the time of exposure to radiation), age is the age of the patient after treatment at the year of interest, V_T is the total volume of the organ, β_{EAR} is the initial slope, and $\mu(ageX, age)$ is the modifying function,

$$\mu(ageX, age) = \exp\left[\gamma_e (ageX - 30) + \gamma_a \ln\left(\frac{age}{30}\right)\right], \quad (10)$$

with γ_e and γ_a being the age modifying parameters.

Eq. (11) gives the risk equivalent dose (RED) mechanistic model which accounts for the effects of cell killing and fractionation

$$RED(D) = \frac{e^{-\alpha'D}}{\alpha'R} \left(1 - 2R + R^2 e^{\alpha'D} - (1-R)^2 e^{\frac{\alpha'R}{1-R}D}\right), \quad (11)$$

where R is the repopulation/repair parameter, α' is given by

$$\alpha' = \alpha + \beta \frac{D_i}{D_T} d_T, \quad (12)$$

and D_T and d_T represent the prescribed dose to the target volume and the corresponding dose per fraction, respectively. The values used for these parameters are listed in Table 7.

The probability of cancer induction for each year after exposure is given by EAR^{org} . Hence it is necessary to convert the probabilities from cumulative to differential.

2.4.4. Transitions between diseased states

In reality, it is possible for more than one injury or SPC to affect a patient at a given time which will in turn affect their quality of life. It is therefore important to account for this in the model. As mentioned in Section 2.1, there is a separate Markov state for every possible combination of injuries and SPCs that can affect a patient at a given time. To transition between any two states that are within the Diseased group, at least one of three possible events E must occur:

- development of an SPC, E_{SPCD} ,
- development of an NTC, E_{NTCD} ,
- recovery from an NTC, E_{NTRC} .

These events were modelled as independent events, and thus it is possible for more than one injury or SPC to develop in a given cycle. For example, the probability of developing injury 1 (E_1) and injury 2 (E_2) in the same cycle is given by

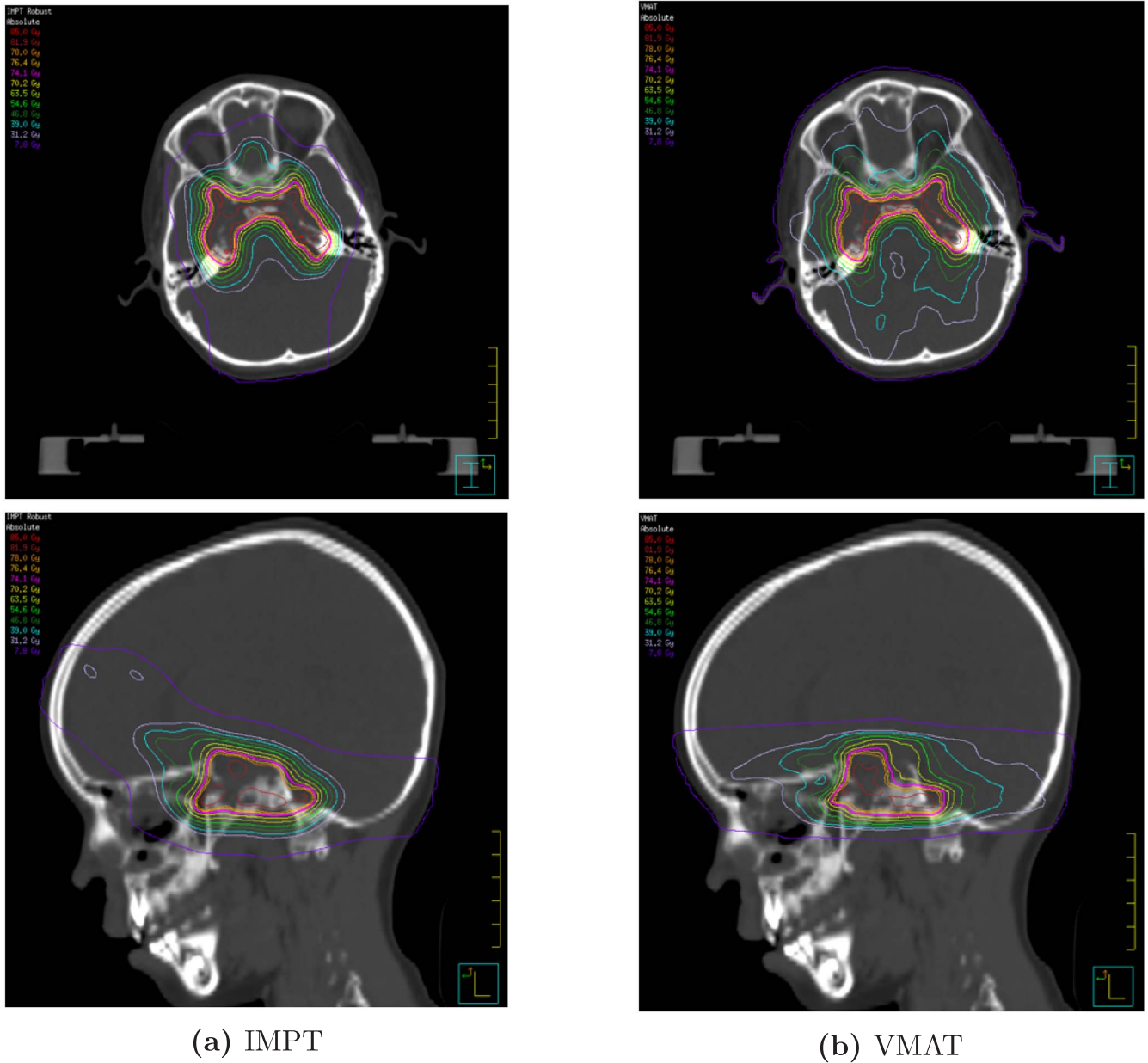


Fig. 2. The CT scan of the patient with the isodose contours for each treatment plan.

$$P(E_1 \cap E_2) = NTCP_1 \times NTCP_2. \tag{13}$$

Note that while it is possible to recover from an injury and develop another in the same cycle, it is not possible to develop and recover from a given injury in the same cycle.

Injuries and SPCs also develop independently of whether the tumour was successfully treated in this model. As a result, the presence or absence of the initial primary cancer was treated as a constant factor and was modelled separately to the transitions between the Diseased states. For example, the probability of transitioning to a Diseased state is the same regardless of whether the treatment was successful.

2.4.5. Death probabilities

The probabilities of a Sim transitioning to the Deceased state are not based on radiobiological models. In this model, there are three ways in which a Sim can transition to the Deceased state:

- Death from the initial primary cancer. This transition probability is time-dependent with a different probability for each year after

treatment. The complementary probability, the probability of surviving cancer, was calculated before converting to death probability. It was assumed that cancer survival followed an exponential distribution. The survival probability was calculated by integrating the exponential density function,

$$f(x, \lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases} \tag{14}$$

with a clinically derived decay parameter λ (listed in Table 8). The function is scaled such that the total time integrated probability of dying of cancer is equal to the area beneath the curve. The trapezoidal rule method as described in Section 2.4.2 was used to integrate the function between two particular years x_1 and x_2 to give a probability of dying from cancer in a particular year.

- Death from injury or second primary cancer. This is also time-dependent and derived using the same method as the primary cancer death probability. Relevant parameters are listed in Table 8. These probabilities are distinct from initial primary cancer death

probabilities as the time of injury or second cancer onset is variable. Therefore, the probability of dying from an injury or second cancer depends on how long it has been affecting a patient. Implementing these presents a practical challenge due to the ‘memoryless’ property of Markov models where the probability of transitioning to a particular state only depends on the current state and not on any previous states. To address this challenge, a different state was used to represent each of the possible time periods that an injury or second cancer had been affecting the patient. For example, for a given injury, there is a separate state for the cases where the injury had been present for one year, for two years, for three years and so on. The appropriate death probability was assigned to each of these states.

- Death from unrelated causes. This probability is based on data from life tables obtained from the Australian Bureau of Statistics (ABS) [19]. The life tables give the probability that a person will die for any given year in their life and hence this is a time-dependent probability. Including the effect of unrelated death results in a more realistic estimate for the life expectancy of the patient.

The model has the ability to distinguish between these different causes of death. Each living state is associated with a probability of dying from each of the various causes listed above. These probabilities are added to the list of possible transitions that can be made from that state. Section 2.3 describes how pseudo-random numbers are used to determine the new state at the end of a cycle given the state at the beginning of the cycle.

2.4.6. Recovery probabilities

The possibility of recovery from an injury once developed is included in the model. The probability of recovering from a given injury is calculated from a normal distribution in a similar fashion to the cancer and injury death probabilities. This normal distribution requires estimates of the mean and standard deviation of the recovery times, along with an overall recovery probability.

3. Clinical example

A 6-year-old male presenting with base of skull chordoma was considered for the purposes of testing the simulation tool. Treatment plans for the delivery of 78 Gy in 39 fractions to the tumour with IMPT (Fig. 2a) and with VMAT (Fig. 2b) were retrospectively generated by using the patient computed tomography (CT) scans and MRI images as input to the research release of Philips Pinnacle³ (Amsterdam, Netherlands) treatment planning software. The treatment plans underwent robust optimisation. In each treatment plan for this patient, each critical structure (healthy tissue) and the target volume corresponding to the tumour were contoured by a clinician. A differential DVH was generated for each of these regions. There was also a DVH for the total volume of normal tissue. The injuries considered are listed in Appendix A. Several injuries were considered for some structures. The TCPs, NTCPs and SPCIPs that were calculated for each treatment plan for this patient are listed in Appendix Tables 10–12, respectively.

3.1. Results

The model predicted the treatment response for a hypothetical cohort of patients, with each being a copy of the patient under consideration. The cohort size was chosen to be $5 \cdot 10^5$, as this gave a standard error of less than 1%. The model returned the median life expectancy of all the Sims along with the median QALE. The results for this clinical example are listed in Table 1. The estimates of the life expectancy (LE) are equal for both treatments but the QALE is greater for the IMPT case.

The life expectancy is related to the cause of death. For the cohort representing this patient, most deaths were not due to the cancer or

Table 1

The estimates of the median raw life expectancies (LE) and the median QALEs in years for each treatment modality.

	IMPT	VMAT
Raw LE (years)	77.5	77.5
QALE (years)	76.5	74.9

Table 2

The proportions of the hypothetical cohort who died as a result of various causes for each treatment modality. Injuries were not considered to be life threatening in this example. Note that the simulation ceases when a Sim reaches 100 years of age.

Cause of death	IMPT	VMAT
Primary cancer	5.98%	8.21%
Second primary cancer	1.36%	1.55%
Unrelated death	91.5%	89.0%
Reached 100 (did not die during the simulation)	1.31%	1.20%

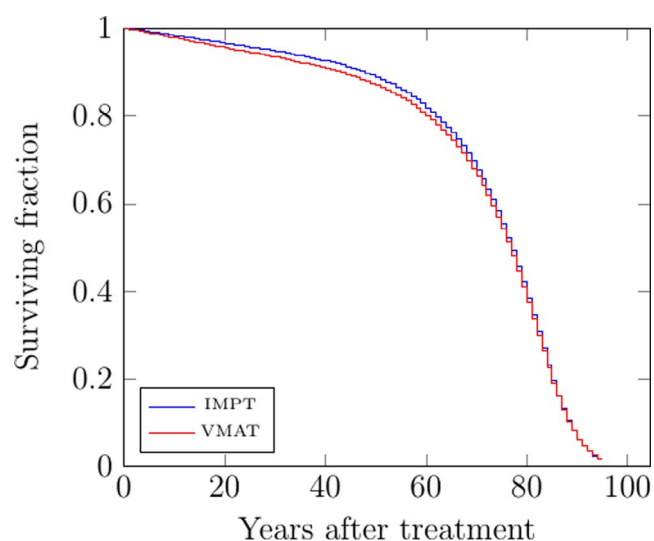


Fig. 3. Kaplan-Meier raw survival curves. A greater survival rate is predicted for IMPT in the earlier years after treatment. The raw data has been binned yearly as it has integer values. The 95% confidence intervals are indicated by shaded areas but are small in magnitude.

treatment. The proportions of the cohort who died as a result of other causes are listed in Table 2.

The Kaplan-Meier survival curves are given in Fig. 3 where the standard deviation used to estimate the 95% confidence intervals ($1.96SD$) was calculated using Greenwood’s formula (15) [20],

$$sd(\hat{S}_{(j)}) = \hat{S}_{(j)} \left[\sum_{k \leq j} \frac{y_k}{n_k(n_k - y_k)} \right]^{1/2}, \quad (15)$$

where $\hat{S}_{(j)}$ is the survival fraction estimate at year j , y_k is the number of deaths that occurred in year k , and n_k is number surviving at year k .

Kaplan-Meier survival curves were also generated for the QALE, and are given in Fig. 4. A consistent difference in the treatments is apparent.

The proportions of Sims suffering an injury or second cancer are listed in Table 3. The model predicted that there would be a small probability of developing a second cancer for both treatments. There is a larger probability that the patient would suffer a radiation induced injury if treated with VMAT, most likely due to the proximity of the tumour to the critical structures.

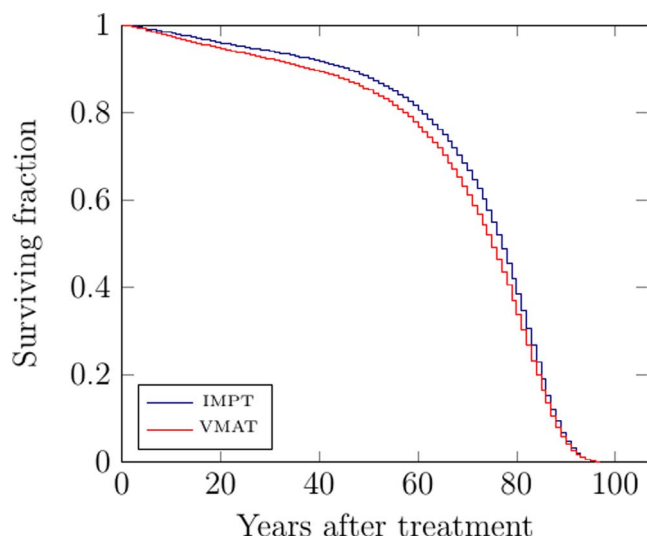


Fig. 4. Kaplan-Meier QALE survival curves. The raw data has been binned yearly. The 95% confidence intervals are indicated by shaded areas but are small in magnitude.

Table 3

The proportion of the cohort that suffered a second cancer or injury.

	IMPT	VMAT
Proportion suffered injury	1.87%	10.3%
Proportion suffered cancer	5.68%	6.04%

3.2. Significance testing

A log-rank test [20] was used to determine the statistical significance of the results. This test allows a non-parametric two-sample comparison of survival data, which is appropriate in this case as not all Sims necessarily reach the Deceased state before they reach an age of 100. The null hypothesis of the log-rank test is that the hazard rate h at a particular year i is equal for two treatments A and B ,

$$H_0(i): h_{Ai} = h_{Bi} \quad (16)$$

where the hazard rate is defined by $h_i = y_i/n_i$, with y_i being the number of deaths and n_i being the number at risk (the number alive at the beginning of year i). Under the null hypothesis, y_i has mean E_i and variance V_i as given by the hypogeometric distribution,

$$\begin{aligned} E_i &= n_{Ai} n_{di} / n_i \\ V_i &= n_{Ai} n_{Bi} n_{di} n_{si} / [n_i^2 (n_i - 1)], \end{aligned} \quad (17)$$

where n_{di} , n_{si} and n_i are the total number of deaths in year i , the total number surviving at the end of year i , and total number at risk at the beginning of year i , respectively, between the two treatments. The log-rank statistic Z is then defined to be

$$Z = \frac{\sum_{i=1}^N (y_{Ai} - E_i)}{(\sum_{i=1}^N V_i)^{1/2}}. \quad (18)$$

By the Central Limit Theorem, under the null hypothesis Z follows a standard normal distribution. A two-tailed log-rank test was used to compare the IMPT and VMAT raw Kaplan-Meier survival curves, as well as the QALE Kaplan-Meier survival curves. A statistically significant difference (p -value < 0.05) between the treatments was detected for both the raw and quality adjusted survival curves.

4. Discussion

A Markov model that predicts the clinical outcome after radiotherapy has been developed and tested with two alternative treatment

plans for an example patient. The results of this demonstration suggest that both the overall survival probability and the quality adjusted survival would be significantly improved with the use of IMPT rather than VMAT. This can be explained by the greater TCP associated with the IMPT plan. In addition, many of the NTCPs associated with the IMPT plan are an order of magnitude smaller than those associated with the VMAT plan (Appendix C). Therefore, this patient could expect to live both longer and with an enhanced quality of life with a reduced probability of suffering a radiation induced injury or second cancer. This is particularly true in the later years after treatment.

The strength of this Markov model is that it allows clear visualisation of the likely patient outcome in terms of not only survival, but also complication risks. Furthermore, it has the ability to distinguish between alternative causes of high injury rates. For example, both treatment plans used in the current example could result in complications related to the parotids due to the elevated dose in this region. These features could be a valuable aid for a clinician faced with the task of prescribing an optimal treatment for a tumour. If the treatment plans used for the demonstration were to be compared using the model of Langendijk et al. [3] which compares NTCPs only, then the result would most likely be similar, as the collective NTCP for VMAT was greater than IMPT. However, the model presented in the current work also accounts for a difference in TCP between treatment plans (see Table 2).

The presented model has several limitations which may have reduced the accuracy of its predictions. The results are predominantly determined by the TCP, NTCP and SPCIP radiobiological models. While efforts were made to include clinically founded parameters in these models and the calculations of other transition probabilities, it was not always possible to obtain relevant estimates in the literature. In addition, while the QALE is an important metric estimated by the model, its value depends strongly on the QoL utilities assigned to each state which were not clinically founded in all cases. In such instances, estimates were used (see Appendix B). In other cases, it was not possible to source entirely appropriate parameters. For example, the probabilities of dying from a radiation induced brainstem glioma was based on adult patient data, yet the model was applied to a paediatric patient in this work. It was also not possible to source a quality of life utility for spinal cord myelitis, and a utility for spinal stenosis was taken as an approximation. However, these values are easily modified in a situation where new studies come to light. Indeed, following the suggestions of Langendijk et al. [3], the input database of model parameters should be continually updated with follow-up data acquired as part of the referral program. NTCP model and dose uncertainties are known to have a significant impact on the accuracy of model-based patient selection [21]. These parameter uncertainties are distinct from the stochastic uncertainty that has been considered in this work. Future development of this model will account for the uncertainties associated with the DVH and radiobiological parameters, allowing for representation of the overall uncertainties in the final results.

The radiobiological models themselves also have limitations. The advantage of the Lyman-Kutcher-Burman NTCP formalism is that it is commonly employed in modelling studies and parameters are therefore widely available. However, only severe radiation induced injuries were considered in this work due to the difficulty in finding parameters for less severe injuries. These types of injuries will be considered in future applications of the model. The SPCIP was based on the Schneider model of radiation-induced cancer [17] which has been developed for dose ranges relevant in radiotherapy. However, that model has limitations. For example, the dose delivered to the whole body can result from a variety of dose distributions in the organs.

Another limitation of the model was that the course of the patient's disease was greatly simplified. For example, the possibility of metastasis of the initial primary cancer or second primary cancers was omitted. A constant quality of life was assumed for these cancer states. Furthermore, it was assumed that if complete control was achieved, there was no possibility of recurrence which also represents a deviation

from reality. These simplifications were made as the treatments used for the initial primary cancer alone were the subjects of comparison. Likewise, any treatments for subsequent injuries or second cancers that would normally be carried out routinely were not considered directly in the model. However, the effects of disease progression and metastasis are important factors in determining both life expectancy and quality of life. For more reliable estimates of these values, disease progression would need to be considered in future versions of the model.

Proton therapy is known to be a significantly more expensive treatment compared with X-ray therapy. As a result, cost-effectiveness can be an important consideration in proton therapy referral. Cost-effectiveness for left-sided breast cancer has been evaluated by Lundkvist et al. [22]. Due to the variability of both treatment costs and healthcare system economics between countries, cost can be a challenging factor to consider. As such, cost-effectiveness has not been considered in the presented model. However, it may be considered in future developments within this framework.

Although the presented model has strengths, its inevitable limitations may reduce its validity. Thorough internal model validation has been carried out where components of the code were checked for consistency. However, performing external model validation where model results are compared with real-world results is a stronger form of validation [23]. Such analysis will be conducted in future studies to

ensure the reliability of the model predictions.

5. Conclusion

A radiobiological-based Markov model for aiding the decision making in proton therapy referral has been presented with the capabilities being demonstrated with a test case. The Markov model allows comparisons to be made between IMPT and X-ray therapy on an individual patient basis. The model suggests that the BOSCh demonstration patient considered in this work would likely receive a significant benefit if treated with IMPT rather than VMAT in terms of a reduced risk of injury. The accuracy of the model is reliant on the quality of the input calculation parameters. The concept of *in silico* clinical trials could be used to gradually refine the accuracy of the input data over time.

Acknowledgements

The authors wish to thank Peter Rhodes for the early development of the project, Alexandre Santos for the development of the SPCIP code and Raymond Dalfsen for producing the VMAT and IMPT patient plans used in this study. The third author acknowledges the support of ACEMS (ARC Centre of Excellence for Mathematical and Statistical Frontiers).

Appendix A. Model parameters

A.1. Tumour control probability

Several studies have been conducted to determine the 5-year local control rate of chordoma for a given median dose using X-ray or proton therapy (without chemotherapy) [24–32]. This data was used to plot a TCP curve. D_{50} was then varied to obtain γ within a 90% confidence interval. The optimum value was yielded by minimising the width of this confidence interval. The results are listed in Table 4 along with the α/β value that was obtained from the literature.

Table 4
Parameters used in the tumour control probability (TCP) calculation with (2).

D_{50} (Gy)	α/β (Gy)	α (Gy^{-1})	γ
60.0	2.45 [33]	0.053	2.06

A.2. Normal tissue complication probabilities

The parameters used in the NTCP calculations are listed in Tables 5 and 6.

Some data listed in Table 6 was derived from histogram data of injury incidences. A normal distribution was fitted to the data once plotted and the best fit mean and standard deviation parameters of the distribution were taken as the estimates listed in Table 6.

Table 5

Parameters used for the calculation of the all-time NTCP for each injury using the LKB model. Burman et al. was used as the source of the parameters unless otherwise stated.

Tissue	α/β (Gy)	Injury/Endpoint	n	m	TD_{50} (Gy)
Brainstem	2.5 [34]	Necrosis	0.16	0.14	65
Spinal cord	0.87 [35]	Myelitis	0.05	0.175	66.5
Ear	3.0 ¹	Acute serous otitis	0.01	0.15	40
		Chronic serous otitis	0.01	0.095	65
Optic nerves and chiasm	3.0 ¹	Blindness	0.25	0.14	65
Parotids	3.0 ¹	Xerostomia	0.70	0.18	46
Lens	3.0 ¹	Cataract	0.30	0.27	18

¹ Assumed as the default 3.0 for late responding tissue.

Table 6

Estimates of the mean (\bar{x}) and standard deviations (σ_x) of the time taken for each injury to develop after treatment. Where it was not possible to find an estimate of the standard deviation, it was assumed to be 6 months.

Tissue	Injury	\bar{x} (months)	σ_x (months)	Source
Brainstem	Necrosis	17	6*	[36]
Spinal cord	Myelitis	13.6	3.9	[37]
Ear	Acute serous otitis	12*	6*	–
	Chronic serous otitis	12*	6*	–
Optic nerves and chiasm	Blindness	18	6*	[16]
Parotids	Xerostomia	0.25	6*	[38]
Lens	Cataract	30	6	[16]

* Not clinically founded.

A.3. Second primary cancer induction probabilities

The parameters used for the SPCIP calculation are listed in Table 7 [17].

Table 7

Parameters used for the calculation of the SPCIP for each year after treatment for each cancer site considered. The same parameters were used for all sites within the brain and central nervous system.

Tissue	β_{EAR}	α	R	γ_c	γ_a
Brain/CNS	0.70	0.018	0.93	–0.024	2.38
Normal tissue	74	0.089	0.17	–0.024	2.38
Parotids	0.73	0.087	0.23	–0.024	2.38

A.4. Death probabilities

The parameters used for the calculation of the death probabilities are listed in Table 8. Where estimates of the mean were given by a source, the decay parameter λ was assumed to be the inverse. Where estimates of the median m were given, it was assumed $\lambda = \ln(2)/m$. Estimates do not assume that any particular treatment is undertaken, although most studies will involve treatment of the patients. The impact on the quality of life of these treatments is not considered in the Markov model.

Table 8

Estimates of the decay parameter, λ , and the total probability of surviving the cancer altogether, $P_{survive}$, that were used to calculate the yearly probabilities of death as a result of cancer.

Tissue	λ years ⁻¹	$P_{survive}$	Source	Comments
Base of skull chordoma	0.9617	0.10	[39]	Derived from 3 and 5 year survival rates after relapse
Brainstem (glioma)	0.154	0.45 (at five years)	[40]	Data for adult glioma
Spinal cord	1.39	0.21	[41]	Survival after surgery for malignant astrocytoma (the most common spinal cord tumour [42] which are themselves generally rare).
Ear	1.0*	0.61	[43]	Carcinoma of external auditory canal and middle ear
Optic nerves and chiasm	0.0526	0.44	[44]	Chiasmatal glioma only considered
Parotids	0.438	0.45	[45]	Overall survival probability including cases where there is and is not local-regional control
Normal tissue	1.0*	0.1*	–	–

* Not clinically founded.

Appendix B. State QoL utilities

The quality of life utilities applied in the Markov model are listed in Table 9. For Diseased states where there is more than one injury or second cancer, the assigned utility is a multiplication of the utilities of the states where there is only one of each injury or second cancer.

The states representing the cases of second primary cancers were all assigned a value of 0.8 in accordance with the Eastern Cooperative Oncology Group (ECOG) performance status [47] as grade 1 (with utility 0.8) gave the most accurate description of these states. With similar reasoning, injuries were assigned a default value of 0.6.

Table 9

Estimates for the QoL utilities for states in the Markov model. The states representing a second primary cancer have default utilities of 0.8. Where it was not possible to find an appropriate QoL utility for a particular injury, a default value of 0.6 was assigned.

State	Utility	Source	Comments
Skull base chordoma	0.72	[10]	
Brainstem necrosis	0.6		Not clinically founded
Spinal cord myelitis	0.7	[46]	Utility for spinal stenosis (confined to manual wheelchair) taken as an approximation for myelitis
Acute serous otitis	0.5	[46]	
Chronic serous otitis	0.7	[46]	
Blindness	0.33	[46]	Complete blindness
Xerostomia	0.826	[10]	
Cataracts	0.6	[46]	Advanced lens opacity

Appendix C. Calculated transition probabilities

The transition probabilities that were derived from radiobiological models are listed in Tables 10–12.

The SPCIP calculation returns a probability for each year after treatment and these values are used as input for the Markov model. They are listed here as combined probabilities for each tissue which were calculated by integrating the function that represents the probability of developing a second cancer for each year after treatment. The combined probabilities are listed in Table 12 for each structure.

Table 10

The values for the TCP that were calculated and used as input for the Markov model for each treatment modality.

	IMPT	VMAT
TCP	0.927	0.899

Table 11

The values for the all-time NTCP (combined NTCPs over all years from treatment to the maximum possible age of 100) that were calculated for each injury using (5) for each treatment modality.

Injury	Tissue	IMPT	VMAT
Brainstem necrosis	Brainstem	$7.97 \cdot 10^{-5}$	$4.36 \cdot 10^{-6}$
Acute serous otitis	Cochlea	$1.11 \cdot 10^{-5}$	$1.52 \cdot 10^{-6}$
Chronic serous otitis	Cochlea	$1.67 \cdot 10^{-16}$	$< 10^{-16}$
Cataracts	Lens	$1.08 \cdot 10^{-4}$	$1.86 \cdot 10^{-4}$
Blindness	Optic nerves and chiasm	$5.20 \cdot 10^{-13}$	$1.03 \cdot 10^{-12}$
Xerostomia	Parotids	$1.18 \cdot 10^{-2}$	0.10
Spinal cord myelitis	Spinal cord	$7.14 \cdot 10^{-3}$	$4.81 \cdot 10^{-3}$

Table 12

The all-time SPCIP values (combined SPCIPs over all years from treatment to the maximum possible age of 100) that were calculated for each treatment modality.

Tissue	IMPT	VMAT
Normal tissue	$6.66 \cdot 10^{-3}$	$1.43 \cdot 10^{-2}$
Brainstem	$1.93 \cdot 10^{-2}$	$1.74 \cdot 10^{-2}$
Cochlea	$2.15 \cdot 10^{-2}$	$2.02 \cdot 10^{-2}$
Optic nerves and chiasm	$4.68 \cdot 10^{-4}$	$3.53 \cdot 10^{-3}$
Parotids	$6.25 \cdot 10^{-3}$	$4.95 \cdot 10^{-3}$
Spinal cord	$4.66 \cdot 10^{-2}$	$5.16 \cdot 10^{-2}$

References

- Armoogum KS, Thorp N. Dosimetric comparison and potential for improved clinical outcomes of paediatric CNS patients treated with protons or IMRT. *Cancers* 2015;7(2):706–22.
- Steneker M, Lomax A, Schneider U. Intensity modulated photon and proton therapy for the treatment of head and neck tumors. *Radiother Oncol* 2006;80(2):263–7.
- Langendijk JA, Lambin P, De Ruyscher D, Widdler J, Bos M, Verheij M. Selection of patients for radiotherapy with protons aiming at reduction of side effects: the model-based approach. *Radiother Oncol* 2013;107(3):267–73.
- Stokkevåg CH, Schneider U, Muren LP, Newhauser W. Radiation-induced cancer risk predictions in proton and heavy ion radiotherapy. *Physica Med* 2017.
- Yoshimura T, Kinoshita R, Onodera S, Toramatsu C, Suzuki R, Ito YM, et al. NTCP modeling analysis of acute hematologic toxicity in whole pelvic radiation therapy for gynecologic malignancies—A dosimetric comparison of IMRT and spot-scanning proton therapy (SSPT). *Physica Med* 2016;32(9):1095–102.
- Magni P, Quaglini S, Marchetti M, Barosi G. Deciding when to intervene: a Markov decision process approach. *Int J Med Inf* 2000;60(3):237–53.
- Hauskrecht M, Fraser H. Planning treatment of ischemic heart disease with partially observable Markov decision processes. *Artif Intell Med* 2000;18(3):221–44.
- Alagoz O, Hsu H, Schaefer AJ, Roberts MS. Markov decision processes: a tool for sequential decision making under uncertainty. *Med Decis Making* 2010;30(4):474–83.
- Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-transition modeling: a report of the ISPOR-SMDM modeling good research practices task force-3. *Value Health* 2012;15(6):812–20.
- Ramaekers BL, Grutters JP, Pijls-Johannesma M, Lambin P, Joore MA, Langendijk JA. Protons in head-and-neck cancer: bridging the gap of evidence. *Int J Radiat Oncol Biol Phys* 2013;85(5):1282–8.
- Schulz S, Guntrum F, Bäumer C, Timmermann B. Expected clinical benefits and challenges of particle therapy for paediatric tumours. *Physica Med: Eur J Med Phys* 2016;32:184–5.
- Allen Li X, Alber M, Deasy JO, Jackson A, Ken Jee K-W, Marks LB, et al. The use and QA of biologically related models for treatment planning: short report of the TG-166 of the therapy physics committee of the AAPM. *Med Phys* 2012;39(3):1386–409.
- Lyman JT. Complication probability as assessed from dose-volume histograms. *Radiat Res* 1985;104(2s):S13–9.
- Kutcher GJ, Burman C. Calculation of complication probability factors for non-uniform normal tissue irradiation: the effective volume method gerald. *Int J Radiat Oncol Biol Phys* 1989;16(6):1623–30.
- Bentzen SM, Constine LS, Deasy JO, Eisbruch A, Jackson A, Marks LB, et al. Quantitative analyses of normal tissue effects in the clinic (QUANTEC): an introduction to the scientific issues. *Int J Radiat Oncol Biol Phys* 2010;76(3):S3–9.
- Rubin P, Constine LS, Marks LB. ALERT: adverse late effects of cancer treatment, volume 2, radiation oncology Springer-Verlag, Berlin Heidelberg; 2014.
- Schneider U, Sumila M, Robotka J. Site-specific dose-response relationships for cancer induction from the combined Japanese A-bomb and Hodgkin cohorts for doses relevant to radiotherapy. *Theor Biol Med Modell* 2011;8(1):1.
- Santos AM, Marcu LG, Wong CM, Bezak E. Risk estimation of second primary cancers after breast radiotherapy. *Acta Oncol* 2016:1–7.
- Life Tables, States, Territories and Australia, 2013–2015; 2016. <http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/3302.0.55.001Main+Features12013-2015?OpenDocument>, accessed: 23-06-2017.
- Efron B, Hastie T. Computer age statistical inference. Cambridge University Press; 2016.
- Bijman RG, Breedveld S, Arts T, Arestreidou E, de Jong MA, Granton PV, et al. Impact of model and dose uncertainty on model-based selection of oropharyngeal cancer patients for proton therapy. *Acta Oncol* 2017:1–7.
- Lundkvist J, Ekman M, Ericsson SR, Isacson U, Jönsson B, Glimelius B. Economic evaluation of proton radiation therapy in the treatment of breast cancer. *Radiother Oncol* 2005;75(2):179–85.
- Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: a report of the ISPOR-SMDM modeling good research practices task force-7. *Med Decis Making* 2012;32(5):733–43.
- Munzenrider JE, Liebsch NJ. Proton therapy for tumors of the skull base. *Strahlenther Onkol* 1999;175(2):57–63.
- Hug EB, Sweeney RA, Nurre PM, Holloway KC, Slater JD, Munzenrider JE. Proton radiotherapy in management of pediatric base of skull tumors. *Int J Radiat Oncol Biol Phys* 2002;52(4):1017–24.
- Noël G, Feuvret L, Calugaru V, Dhermain F, Mammari H, Haie-Méder C, et al. Chordomas of the base of the skull and upper cervical spine. one hundred patients irradiated by a 3D conformal technique combining photon and proton beams. *Acta Oncol* 2005;44(7):700–8.
- Ares C, Hug EB, Lomax AJ, Bolsi A, Timmermann B, Rutz HP, et al. Effectiveness and safety of spot scanning proton radiation therapy for chordomas and chondrosarcomas of the skull base: first long-term report. *Int J Radiat Oncol Biol Phys* 2009;75(4):1111–8.
- Rombi B, Ares C, Hug EB, Schneider R, Goitein G, Staab A, et al. Spot-scanning proton radiation therapy for pediatric chordoma and chondrosarcoma: clinical outcome of 26 patients treated at Paul Scherrer Institute. *Int J Radiat Oncol Biol Phys* 2013;86(3):578–84.
- Yasuda M, Bresson D, Chibbaro S, Cornelius JF, Polivka M, Feuvret L, et al. Chordomas of the skull base and cervical spine: clinical outcomes associated with a multimodal surgical resection combined with proton-beam radiation in 40 patients. *Neurosurg Rev* 2012;35(2):171–83.
- Staab A, Rutz HP, Ares C, Timmermann B, Schneider R, Bolsi A, et al. Spot-scanning-based proton therapy for extracranial chordoma. *Int J Radiat Oncol Biol Phys* 2011;81(4):e489–96.
- Cummings BJ, Hodson DI, Bush RS. Chordoma: the results of megavoltage radiation therapy. *Int J Radiat Oncol Biol Phys* 1983;9(5):633–42.
- Fuller DB, Bloom JG. Radiotherapy for chordoma. *Int J Radiat Oncol Biol Phys* 1988;15(2):331–9.
- Henderson FC, McCool K, Seigle J, Jean W, Harter W, Gagnon GJ. Treatment of chordomas with cyberknife: Georgetown University experience and treatment recommendations. *Neurosurgery* 2009;64(2):A44–53.
- Mayo C, Yorke E, Merchant TE. Radiation associated brainstem injury. *Int J Radiat Oncol Biol Phys* 2010;76(3):S36–41.
- Kirkpatrick JP, van der Kogel AJ, Schultheiss TE. Radiation dose-volume effects in the spinal cord. *Int J Radiat Oncol Biol Phys* 2010;76(3):S42–9.
- Debus J, Hug E, Liebsch N, O'farrel D, Finkelstein D, Efid J, et al. Brainstem tolerance to conformal radiotherapy of skull base tumors. *Int J Radiat Oncol Biol Phys* 1997;39(5):967–75.
- Abbatucci J, Delozier T, Quint R, Roussel A, Brune D. Radiation myelopathy of the cervical spinal cord: time, dose and volume factors. *Int J Radiat Oncol Biol Phys* 1978;4(3–4):239–48.
- Deasy JO, Moiseenko V, Marks L, Chao KC, Nam J, Eisbruch A. Radiotherapy dose-volume effects on salivary gland function. *Int J Radiat Oncol Biol Phys* 2010;76(3):S58–63.
- Fagundes MA, Hug EB, Liebsch NJ, Daly W, Efid J, Munzenrider JE. Radiation therapy for chordomas of the base of skull and cervical spine: patterns of failure and outcome after relapse. *Int J Radiat Oncol Biol Phys* 1995;33(3):579–84.
- Landolfi JC, Thaler HT, DeAngelis LM. Adult brainstem gliomas. *Neurology* 1998;51(4):1136–9.
- Cohen AR, Wisoff JH, Allen JC, Epstein F. Malignant astrocytomas of the spinal cord. *J Neurosurg* 1989;70(1):50–4.
- Baleriaux D. Spinal cord tumors. *Eur Radiol* 1999;9(7):1252–8.
- Pfeundner L, Schwager K, Willner J, Baier K, Bratengeier K, Brunner FX, Flentje M. Carcinoma of the external auditory canal and middle ear. *Int J Radiat Oncol Biol Phys* 1999;44(4):777–88.
- Rush JA, Younge BR, Campbell RJ, MacCarty CS. Optic glioma: long-term follow-up of 85 histopathologically verified cases. *Ophthalmology* 1982;89(11):1213–9.
- Storey MR, Garden AS, Morrison WH, Eicher SA, Schechter NR, Ang KK. Postoperative radiotherapy for malignant tumors of the submandibular gland. *Int J Radiat Oncol Biol Phys* 2001;51(4):952–8.
- Tengs TO, Wallace A. One thousand health-related quality-of-life estimates. *Med Care* 2000;583–637.
- Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol* 1982;5(6):649–56.

3.3 Markov model details

In this section the details of the Markov model are described, with a focus on technical aspects that were considered too detailed for the publication. The code is written in the C programming language, which allows for an improved computation time compared with higher-level programming languages such as Python or MATLAB. The program has been written in general terms and, as a result, its strength is its ability to consider any cancer site, with an arbitrary number of treatments and injuries.

The Markov model consists of several phases. These are:

1. Determining the number of Markov states based on the model input,
2. Calculating the transition probabilities using the patient data,
3. Constructing the transition matrix using the transition probabilities,
4. Evaluating the Markov model using a Monte Carlo simulation,
5. Analysing the results.

Each of these stages is explained in more detail in this section.

3.3.1 Determining the number of Markov states

In reality, it is possible for more than one injury (or normal tissue complication (NTC)) or second primary cancer (SPC) to affect a patient at a given time, which will in turn affect their quality of life. It is therefore important to account for this in the model. To model this accurately, a Markov state was allocated for every possible combination of injuries and cancers (including SPCs and the initial primary cancer) that can affect a patient at a given time. The number of Markov states in the Unwell group is a variable which depends on the number of second primary cancers (SPCs), N , and of injuries, M , that are being considered in the model. Therefore, the total number of states, n , is given by

$$n = 2^{M+N+1} + M + N + 2, \quad (3.1)$$

where addition of 1 in the exponent represents the initial primary cancer (the presence of this is independent of the development of injuries and SPCs). It is possible to die from

each injury and cancer (including the initial primary) as well as background death. The additional states signified by the sum $(M + N + 2)$ represent the Deceased states. While it is possible to model death with only one Deceased state, having one for each possible cause of death allows the cause of death to be tracked by the model.

3.3.2 Transition probability calculation

The transition probabilities in the Markov model can be broadly classified as *dose-dependent* and *dose-independent*. The dose-dependent transition probabilities are calculated using input dose-volume histograms (DVHs) for each organ being considered. These histograms give the volume of an organ that receives a given dose and hence are unique to a given patient. The dose-dependent transition probabilities include the tumour control probability (TCP), normal tissue complication probabilities (NTCP) and second primary cancer induction probabilities (SPCIP). The dose-independent transition probabilities include death and injury recovery, and are assumed to be representative of all patients. Another assumption of the Markov model is that it is not possible to recover from a cancer (including both the initial primary cancer and any second cancers). In this section, additional details are provided on select groups of transition probabilities.

3.3.2.1 Normal tissue complication probability (NTCP)

The model has the flexibility to incorporate any number of injuries in the analysis. The NTCP is calculated for each injury and for each potential year after treatment using the methods described in the publication P1, Section 2.4.2. To complete the construction of the normal distribution, which has a user-defined mean and standard deviation, the NTCP for each year is scaled such that the accumulated NTCP from the treatment time to the time at which the patient is 100 years of age is equal to the all-time NTCP calculated with the LKB model [60, 61]. This is done by a recursive adjustment of a linear scaling parameter, as the total response to such a change is non-linear and cannot easily be determined analytically. At each iteration, the all-time NTCP calculated with the LKB is compared with the all-time NTCP corresponding to the normal distribution with

the given mean and standard deviation. Each time point in the distribution is adjusted at each iteration to achieve closer agreement with the NTCP calculated with the LKB model. This process is repeated for each injury, resulting in an NTCP for each possible injury for each year after treatment.

3.3.2.2 Second primary cancer induction probability (SPCIP)

The calculation of the SPCIPs are similar to the NTCPs, in that there is a probability for each year and for each tissue. The difference is that these time-dependent probabilities do not follow a normal distribution, as detailed in the publication P1, Section 2.4.3. In addition, the units of EAR^{org} are per 10^4 person-years and hence EAR^{org} must be divided by 10^4 .

3.3.2.3 Cancer and injury death

A key feature of first-order Markov models is the *memoryless* property, whereby the transition probabilities depend only on the current state and not on any previous states. It was assumed that in reality the death probability changes depending on if and when a cancer or injury develops, which can be represented as a Markov model of arbitrary order. Injury death (and also injury recovery) was assumed to follow a normal distribution, with injury-specific parameters specified by the user. Cancer death was assumed to follow an exponential distribution (Section 2.4.5 of the publication P1). Background death, which is not discussed in this section, depends on the age of the patient at a given time.

It is possible to accurately implement a time-dependent death probability using a first-order Markov chain, by having a separate state for each time period that an injury or cancer could be affecting a patient, and assigning the probability of dying depending on which state the patient is in. The disadvantage of this approach is that the number of Markov states becomes very large, also known as “state explosion”, which decreases computational efficiency as this is a function of the number of Markov states.

To avoid state explosion, an alternative method for modelling time dependence is used in the code. An array is used to track the year in which a particular injury or second

cancer develops. In subsequent years, the death probability that is applied corresponds to the year since it was developed. Injury recovery is also treated in a similar manner, with the yearly recovery probability being applied rather than the death probability. It is assumed that the initial primary cancer, if present, has been present since the initial Markov cycle and hence it is not necessary to track the time since its development.

A flag is also used to track whether an injury has been recovered from. If the flag evaluates to TRUE, the relevant NTCP is set to zero for all subsequent years to prevent the injury from redeveloping.

3.3.2.4 Scaling of death probabilities

For a given state, the true probability of death can be difficult to determine exactly, as it is possible for a patient to die as result of multiple different causes. For example, if a patient is in the Well state, then it is possible to die as a result of unrelated causes only. However, if the patient is in a state representing a cancer, then it is possible to die from either that cancer or unrelated causes. Dying as a result of the cancer or from unrelated causes are not independent events as it is not possible to die from both. Furthermore, the possibility of both a cancer and an injury also exists in this model, and in this case, there would be three possible ways to move to the Deceased state.

In this example, let the event of dying from cancer be Die_C and the event of background death be Die_B , and assume that it is not possible to die from an injury. As these two events are disjoint and exhaustive, the probability of not dying is $1 - Pr(Die_B) - Pr(Die_C)$. Ideally, studies would report the fraction of their study population that died of various causes, which would allow true probabilities to be known, as $Pr(Die_B)$ and $Pr(Die_C)$ would be derived from the same source. However, mortality rates due to various causes are often convoluted in a single reported mortality. This situation becomes increasingly complex with an increasing number of injuries and cancers. As a result, it was necessary in this work to source probabilities from different studies, with the consequence being that the sum of individual death probabilities can exceed 1, which should not be possible. It is not a reasonable approximation in this case to truncate the combined death probability

to 1 and hence assume that death occurs with certainty. Hence, a challenge exists in combining probabilities from multiple sources in a meaningful way in the Markov model.

As an alternative, it was assumed that the probability of moving from a given state to the Deceased state was the maximum of all possible death probabilities:

$$Pr(\text{Die}) = \max_i \{Pr(\text{Die}_i)\}, \quad (3.2)$$

where $Pr(\text{Die}_i)$ is the input probability of dying due to cause i . Clearly, if an injury or cancer is not present in a given state, then it is not possible to die as a result of it.

In order to maintain a distinction between alternative causes of death, all death probabilities are scaled such that their sum is equal to the maximum death probability (before scaling), as given by (3.2). To achieve this, each unscaled death probability is divided by the scaling factor, c , given by

$$c = \frac{\sum_{i=0}^{M+N+2} Pr(\text{Die}_i)}{Pr(\text{Die})}. \quad (3.3)$$

The limitation of this approach is that the true estimates are not preserved. Furthermore, a patient with one potentially fatal injury would have the same probability of dying as a patient with two potentially fatal injuries. However, this is a reasonable assumption to make as often there is a single dominant death probability. This is particularly true when the patient is older and the background death probability becomes large. Alternatively, there is a higher probability of dying in the years immediately after treatment if the treatment was unsuccessful. For the example presented in the publication P1, an alternative method was used that involved scaling all values of the row of the transition matrix such that all values summed to 1, rather than scaling the death probabilities separately. If the scaling method outlined in this section is applied instead, then the estimates of the QALEs are 77.0 for protons and 76.0 for photons. This represents a small difference of 0.5 and 1.1 years, respectively. Therefore, implementing this method of scaling has a small impact on the results, while being more mathematically correct in that transition probabilities to non-deceased states are preserved.

It is also worth noting that the death probabilities are dose-independent and are included to allow for a more realistic estimation of the life expectancy resulting from a given treatment plan. In contrast to the dose-dependent transition probabilities that will differ between the treatment plans being compared, the death probabilities are constant between treatment plans.

3.3.3 The transition matrix

The development of the transition matrix is a complex process due to the large number of states involved. The transition matrix P has dimensions of $n \times n$ where n is the number of Markov states. Each element p_{ij} , $i = 1, \dots, n$, $j = 1, \dots, n$, represents the probability of transitioning from state i to state j .

To transition between any of the states that are within the Unwell group (all Markov states excluding Well and Deceased), at least one of three possible events must occur: development of an SPC, development of an injury, or recovery from an injury. In addition, death must not occur if a transition is made from one Unwell state to another. Injuries and SPCs develop independently of whether the tumour was successfully treated in this model. As a result, the presence or absence of the initial primary cancer was treated as a constant factor and was modelled separately to the transitions between the Unwell states.

Depending on the number of iterations chosen (the number of Sims) and the number of organs at risk (which determines the number of states, see Section 3.3.1), the Monte Carlo model can have a significant computation time. However, as each Sim (as defined in the publication P1, Section 2.3) in the hypothetical cohort is considered separately, it is only necessary to calculate the row of the transition matrix corresponding to the starting state of the Sim at a given cycle. The process of constructing each row of the matrix includes:

1. Efficiently representing the states.
2. Determining which injuries or second cancers contribute to a given transition probability.

3. Multiplying the probabilities of independent events together that all must occur for a given transition.

3.3.3.1 Representing states

Consider the situation where there are M possible injuries and N possible SPCs that could affect the Sim at any given time. When constructing the transition matrix, it is important to be able to efficiently determine which of the states correspond to injured states or cancer states. A method was developed to ensure all combinations were accounted for and that the meaning of each state could be easily specified to allow appropriate quality of life utilities to be applied. A Boolean vector \mathbf{S}_i with $M + N$ elements was used to track which injuries and SPCs are affecting the Sim at the beginning of a particular Markov cycle in the initial state i :

$$\mathbf{S}_{i,\ell} = \begin{cases} 1 & \text{if Sim has NTC/SPC } \ell, \\ 0 & \text{otherwise.} \end{cases} \quad (3.4)$$

Therefore, \mathbf{S}_i represents the state being transitioned from in a particular cycle. Another Boolean vector, \mathbf{S}_j with $M + N$ elements, is used to track which injuries and SPCs are affecting the Sim at the end of a Markov cycle in the state that is transitioned to. The first M elements of \mathbf{S}_i or of \mathbf{S}_j correspond to the injuries and the remaining elements represent the SPCs.

For example, if $M = 3$ and $N = 2$, then the state where the Sim has injury 1, injury 3, and SPC 2 is represented by $\mathbf{S} = [1 \ 0 \ 1 \ 0 \ 1]$. In the Well state, there are no injuries or cancers present and therefore it is represented by $\mathbf{S} = [0 \ 0 \ 0 \ 0 \ 0]$. A vector is not required to represent Deceased states as there is no possibility of leaving absorbing states, and it is also unnecessary to multiply the death probabilities to obtain the final probability of moving to a deceased state, as explained later in Section 3.3.3.2.

It is not necessary to consider the states where both injuries/second cancers and primary cancer are present, as the presence or absence of the primary cancer does not directly affect the transition probabilities between Unwell states. A separate flag is used to track the

presence or absence of the primary cancer, which is constant throughout the simulation. This flag allows the correct probabilities and utilities to be applied at the appropriate year. Hence, the state representing the situation where only the initial primary cancer is present is also represented by $\mathbf{S} = [0 \ 0 \ 0 \ 0 \ 0]$ if $M + N = 5$.

3.3.3.2 Calculation of transition probabilities

The probability of transitioning to the Deceased state for a given starting state is determined using the method outlined in Section 3.3.2.4. To transition between all other Markov states, one or more events must occur. For example, one or more injuries may develop while others do not and others are recovered from. The development of, and recovery from, different injuries and SPCs were modelled as independent events. The development of a second cancer and an injury in the same tissue were also assumed to be independent events, as there are very different cellular processes involved. In this model, it is therefore possible for more than one injury or SPC to develop in a given cycle. As these events are assumed to be independent, the probability of multiple events occurring is given by the multiplication of the probabilities of the events occurring individually. It should be noted that while it is possible to recover from an injury and develop another in the same cycle, it is not possible to develop and recover from the same injury in a given cycle.

There is an event E_ℓ for each second cancer and injury ℓ , $\ell = 1, \dots, M + N$, that is being considered for a given patient (recall that it is not possible to develop or recover from the initial primary cancer once the Markov process has begun and hence it is not necessary to account for this). Therefore, the probability of transitioning from state i to state j at year x after treatment is given by:

$$p_{ij}(x) = (1 - Pr(\text{Die}, i, x)) \prod_{\ell=1}^{M+N} Pr(E_\ell, i, j, x); \quad i = 1, \dots, n - n_d, \quad j = 1, \dots, n - n_d, \quad (3.5)$$

where $n_d = M - N - 2$ is the number of Deceased states, $Pr(E_\ell, i, j, x)$ is the probability of event E_ℓ occurring in year x in the transition from state i to state j , and $Pr(\text{Die}, i, x)$ is the probability of transitioning from state i to the Deceased state. The probabilities

$Pr(E_\ell, i, j, x)$ are determined by comparing the Boolean vector, \mathbf{S}_i , representing the state being transitioned from, i , to the vector, \mathbf{S}_j , representing the state being transitioned to, j . The values of each of the ℓ elements of \mathbf{S}_i and \mathbf{S}_j are determined by the events E_ℓ that occur in the transition. The events include injury development, second cancer development, or injury recovery. The probability of not developing an injury or cancer, or not recovering from an injury is given by the complement of the respective development or recovery probabilities. The event probabilities are given by:

$$Pr(E_\ell, i, j, x) = \begin{cases} \text{NTCP}_\ell(x) & \text{if } \mathbf{S}_{i,\ell} = 0, \quad \mathbf{S}_{j,\ell} = 1, \quad \ell \leq M, \\ \text{SPCIP}_{\ell-M}(x) & \text{if } \mathbf{S}_{i,\ell} = 0, \quad \mathbf{S}_{j,\ell} = 1, \quad M < \ell \leq M + N, \\ 1 - \text{NTCP}_\ell(x) & \text{if } \mathbf{S}_{i,\ell} = 0, \quad \mathbf{S}_{j,\ell} = 0, \quad \ell \leq M, \\ 1 - \text{SPCIP}_{\ell-M}(x) & \text{if } \mathbf{S}_{i,\ell} = 0, \quad \mathbf{S}_{j,\ell} = 0, \quad M < \ell \leq M + N, \\ Pr(\text{IR}_\ell, x) & \text{if } \mathbf{S}_{i,\ell} = 1, \quad \mathbf{S}_{j,\ell} = 0, \quad \ell \leq M, \\ 1 - Pr(\text{IR}_\ell, x) & \text{if } \mathbf{S}_{i,\ell} = 1, \quad \mathbf{S}_{j,\ell} = 1, \quad \ell \leq M, \\ 0 & \text{if } \mathbf{S}_{i,\ell} = 1, \quad \mathbf{S}_{j,\ell} = 0, \quad M < \ell \leq M + N, \\ 1 & \text{if } \mathbf{S}_{i,\ell} = 1, \quad \mathbf{S}_{j,\ell} = 1, \quad M < \ell \leq M + N. \end{cases} \quad (3.6)$$

where $Pr(\text{IR}_\ell, x)$ is the recovery probability for injury ℓ at year x . As the first M elements of \mathbf{S} correspond to injuries, it is necessary to subtract M from the SPCIP index to ensure that the correct probability is applied. For example, if $M = 2$, $N = 3$ and $\ell = 3$, then the relevant SPCIP is SPCIP_1 , even though $\ell = 3$. Note that it is not possible to recover from second cancers. It is necessary to multiply all transition probabilities between living states by the complement of the death probability as it is assumed that death must not occur for a transition to be made to a living state.

The elements p_{ij} , where $1 \leq i \leq n - n_d$ and $n - n_d < j \leq n$, correspond to the death probabilities, which include the probabilities of dying as a result of each injury ($j = n - n_d + 1, \dots, n - n_d + M$), followed by each second cancer ($j = n - n_d + M + 1, \dots, n - n_d + M + N$), the primary cancer, $p_{i,n-1}$, and the probability of dying due to unrelated causes, p_{in} . If an injury or cancer is not present in the state being considered, then the corresponding death probability is set to zero.

The elements p_{ij} , where $n - n_d < i \leq n$ and $1 \leq j \leq n - n_d$, are all zero, as each of these represent the probability of returning from the dead. Furthermore, for $i > n - n_d$ and $j > n - n_d$, if $i \neq j$, then $p_{ij} = 0$ (the probability of moving between deceased states), and if $i = j$ then $p_{ij} = 1$ (the probability of remaining in a Deceased state).

In this model, not all rows of the transition matrix are defined simultaneously. Only the row that represents the state at the beginning of a given cycle is defined. However, to demonstrate (3.6), an example transition matrix for a given year is provided below for the case where there are two injuries only. The states are ordered from left to right (and top to bottom) as follows: Well, Injury 1, Injury 2, Injury 1 and 2, Deceased (due to Injury 1), Deceased (due to Injury 2), and Deceased (due to unrelated causes). Each death probability is scaled using the method described in Section 3.3.2.4, to give the scaled probability of death due to unrelated causes, $Pr(D_b)$ (denoted by D_b), and the scaled probability of death due to injuries 1 and 2, $Pr(D_{i1})$ and $Pr(D_{i2})$, respectively, such that $Pr(D_b) + Pr(D_{i1}) + Pr(D_{i2}) = \alpha$, the maximum of all the possible death probabilities. The event $NTCP_x$ is denoted by I_x for simplicity. The recovery probability is denoted by R_x for injury x . The transition matrix \mathbf{P} is given by

$$\mathbf{P} = \begin{pmatrix} p_{11} & I_1(1 - I_2)(1 - \alpha) & I_2(1 - I_1)(1 - \alpha) & I_1 I_2(1 - \alpha) & 0 & 0 & D_b \\ R_1(1 - I_2)(1 - \alpha) & p_{22} & R_1 I_2(1 - \alpha) & I_2(1 - R_1)(1 - \alpha) & D_{i1} & 0 & D_b \\ R_2(1 - I_1)(1 - \alpha) & I_1 R_2(1 - \alpha) & p_{33} & I_1(1 - R_2)(1 - \alpha) & 0 & D_{i2} & D_b \\ R_1 R_2(1 - \alpha) & R_2(1 - R_1)(1 - \alpha) & R_1(1 - R_2)(1 - \alpha) & p_{44} & D_{i1} & D_{i2} & D_b \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

where the diagonal elements are

$$\begin{cases} p_{11} = (1 - I_1)(1 - I_2)(1 - \alpha), \\ p_{22} = (1 - R_1)(1 - I_2)(1 - \alpha), \\ p_{33} = (1 - I_1)(1 - R_2)(1 - \alpha), \\ p_{44} = (1 - R_1)(1 - R_2)(1 - \alpha). \end{cases} \quad (3.7)$$

All possible transitions are both mutually exclusive and collectively exhaustive. In other words, either a single transition is made in a given Markov cycle or the patient stays in the same state, with the state at the beginning of the cycle being the same as the state at the end of the cycle. The implication is that each row of the transition matrix will sum to 1.

3.3.4 Monte Carlo simulation

The Monte Carlo simulation involves using random numbers to determine which transitions are made in the Markov model. A random number $\eta \in [0, 1]$ can be used for this purpose. If this interval is partitioned into subintervals that represent the magnitude of each of the transition probabilities, then the most likely transition will be the one with the largest magnitude. The interval $[0, 1]$ is partitioned by converting the transition probabilities in a given row of the transition matrix from absolute to cumulative. The result is that $[0, 1]$ contains n subintervals where n is the number of states. This process is demonstrated by Figure 3.1 with three states considered for simplicity.

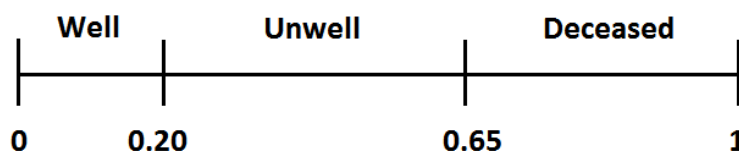


FIGURE 3.1: Demonstration of the conversion of the transition probabilities from absolute to cumulative. The interval $[0, 1]$ is partitioned to represent the magnitudes of each probability.

In the example depicted in Figure 3.1, the row of the transition matrix is $\mathbf{P}_i = [0.2, 0.45, 0.35]$, and the cumulative vector is $\mathbf{C}_i = [0.2, 0.65, 1]$. In general, each element k of \mathbf{C}_i is calculated as:

$$\mathbf{C}_i(k) = \sum_{j=1}^k p_{ij}, \quad \text{for } 1 \leq k \leq n. \quad (3.8)$$

Thus, $\mathbf{C}_i(n) = 1$. If a particular transition is not allowed, for example to state k , then it will have zero width on the interval $[0, 1]$ and $\mathbf{C}_i(k) = \mathbf{C}_i(k - 1)$.

As explained in the publication P1, Section 2.3, if $\mathbf{C}_i(k) < \eta \leq \mathbf{C}_i(k+1)$ then the Sim moves to state k . If $\eta \leq \mathbf{C}_i(1)$, which is the probability of being in the Well state at the end of the cycle (see the definition of \mathbf{P} in Section 3.3.3.2), then the Sim transitions to the Well state.

Once a transition from state i to state j has been made, each $\mathbf{S}_{i,\ell}$ is reset to the values of $\mathbf{S}_{j,\ell}$ for all ℓ , and the process of recalculating the transition probabilities to the allowed states is repeated in the subsequent cycle of the Markov model.

For each cycle, the age and state variables are updated. When a Deceased state is reached, the simulation is terminated.

3.3.5 Analysis of model output

Once the Markov chain has been simulated for each Sim in the cohort, with the life expectancy and QALE determined for each, it is possible to obtain mean life expectancies and QALEs. The mean QALE in particular represents the most important output of the model.

There is a notable technicality in the code, in that the counter for the loop that iterates through each year of the Sim's life starts at zero and hence its value is zero during the first year of the Markov cycle. However, if a Sim were to die in the first year then their life expectancy is 1.

Kaplan-Meier curves, both raw and quality-adjusted are also produced using the data for each Sim. Examples of these are shown in Figures 3 and 4 of the publication P1. Log-rank tests were used to determine the statistical significance of the difference in the survival curves between protons and photons. The results are presented in Table 3.1. A statistically significant difference was found for both the raw and quality-adjusted survival curves.

Along with the survival curves, the fraction of the cohort with any injury or second cancer for each year after treatment is also stored.

	Raw LE	QALE
Z-statistic	8.360	49.143
<i>p</i> -value	<0.0001	<0.0001

TABLE 3.1: Results of the two-tailed log-rank tests comparing raw and quality-adjusted Kaplan-Meier survival curves. The Z-statistic and corresponding *p*-values are shown.

3.4 Model verification

Once the model had been developed, tests were carried out to ensure that it was behaving as expected. These tests included modelling situations in which a patient receives zero dose for the cases with and without an initial tumour present. The tests were used to confirm that the cancer death model and life table data were being processed correctly in the Markov framework. A cohort size of 10^5 was chosen for the purposes of testing the model, as this gave a statistical standard error of less than 1%. The proton treatment plan data used for the publication P1 was used as input for the verification. For the purposes of this testing, the cohort membership was counted at the beginning of each cycle.

3.4.1 Zero dose, tumour absent case

For the case where the tumour was absent, and no dose is received, the expected fraction of Sims surviving at the beginning of year t , for $t \geq 1$ is given by

$$F_t = \prod_{i=1}^{t-1} (1 - Pr_i), \quad (3.9)$$

where Pr_i is the probability of a healthy individual dying in the i^{th} year of their life as given by the life table data. Note that Death is not possible at $t = 0$, the beginning of the first Markov cycle, and hence $Pr_0 = 0$ which gives $F_0 = 1$. The results are given in Figure 3.2.

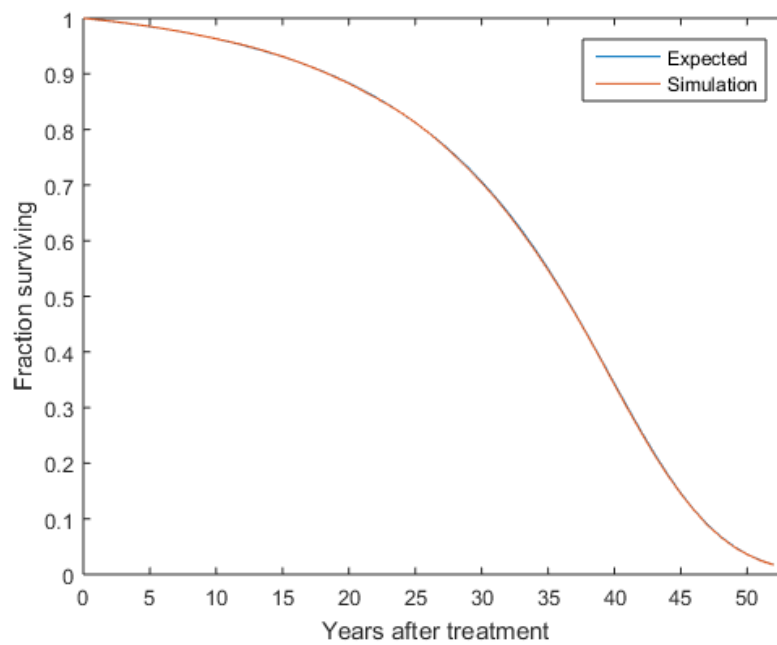


FIGURE 3.2: The surviving fraction of the cohort for each year after treatment, where there is no dose received by any organs and a tumour was absent. As a result, the only enabled death possibility was background death. A good agreement between the output and expectations is apparent.

3.4.2 Zero dose, tumour present case

The expected surviving fraction was also calculated for the case where the tumour was present using (3.9), with Pr_i instead being the time-dependent probability of dying from the primary cancer during a given year i after the treatment. The effect of background death was omitted for the purposes of testing this component of the model. The results of this test are summarised in Figure 3.3.

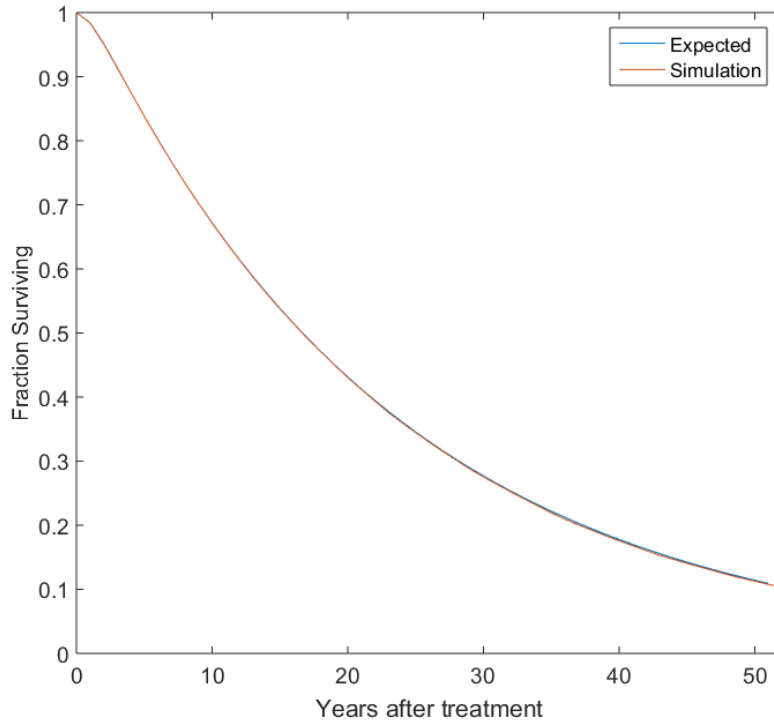


FIGURE 3.3: The surviving fraction of the cohort for each year after treatment, where there is no dose received by any organs and a tumour was present (TCP manually overridden and set to 0). The integral death probability was set to 0.9 and the decay constant needed for the calculation of yearly death probabilities was set to 0.96. Background death was also disabled. OARs outside of the tumour volume received no dose for this test. A good agreement between the output and expectations is apparent.

3.4.3 Uniform dose, tumour absent case

The radiobiological components of the model were also verified. The NTCP and SPCIP components were tested for the case where a tissue received a uniform dose of 50 Gy in 30 fractions, which resulted in a 21% injury occurrence rate.

Figure 3.4 shows a comparison of the expected fraction of the cohort with the injury and fraction predicted by the model. The expected fraction with an injury at the beginning of year t was estimated with:

$$I_t = 1 - [(1 - I_{t-1})(1 - NTCP_{t-1}) + (Pr_{IR})_{t-1}I_{t-1}], \quad \text{for } 1 \leq t < 100, \quad (3.10)$$

where Pr_{IR} represents the probability of injury recovery. The sum on the right-hand-side of (3.10) calculates the fraction without an injury, with the first summand calculating the

fraction without an injury who do not develop one in that cycle and the second calculating the fraction with an injury who recover during the cycle. It is not possible to develop an injury at the beginning of the first cycle.

The injury death model was also tested. The results are presented in Figure 3.5. The expected fraction with an injury at the beginning of year t was estimated using (3.9) with P_i being set to the injury death probability for year i . In this test, the probability of death was set as a constant of 0.01 for all years. Note that Death is not possible at $t = 0$, the beginning of the first Markov cycle, and hence $P_0 = 0$ which gives $S_0 = 1$.

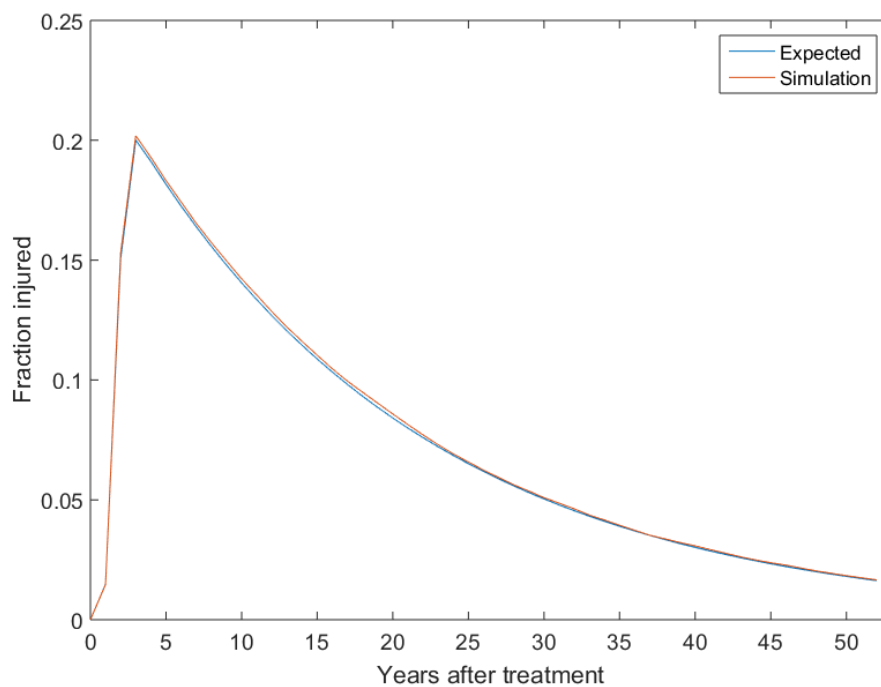


FIGURE 3.4: The fraction of the cohort with an injury, for each year after treatment. The NCTP was set to 0.21 (as a result of a uniform dose delivered to an OAR) for the first year after treatment and 0 for all subsequent years and the probability of injury recovery was set to 0.05. The possibility of death from all causes was disabled for the purposes of this test. A good agreement between the expectations and the output is apparent.

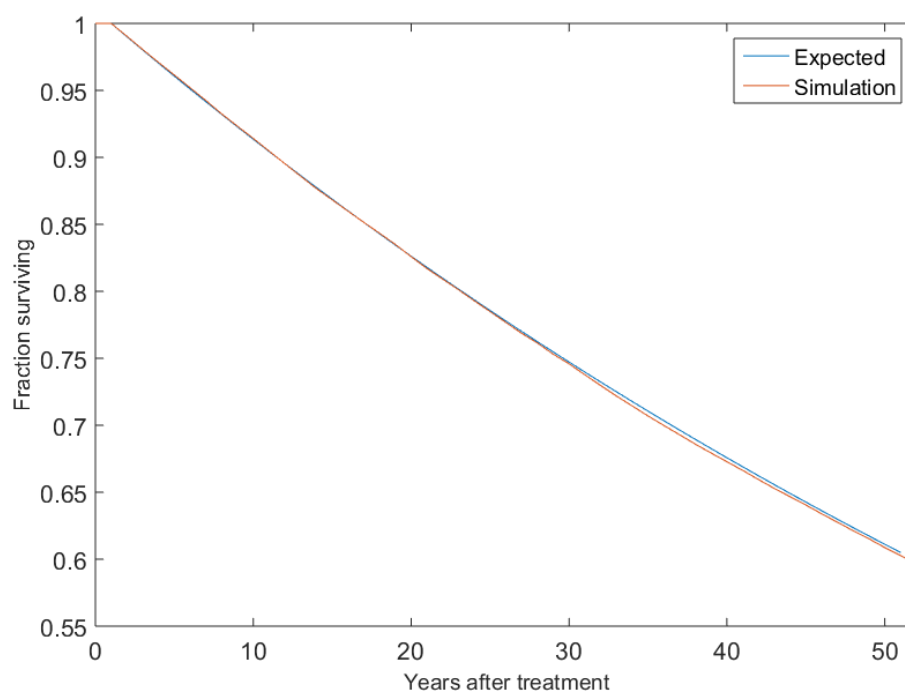


FIGURE 3.5: The surviving fraction for each year after treatment, where an OAR received a uniform dose (resulting in an NTCP of 1.0 for the first year after treatment). The possibility of injury recovery was disabled for the purposes of this test. The injury death probability was set to 0.01, and all other death possibilities were disabled. A good agreement between the expectations and the output is apparent.

3.4.4 Statistical significance

To determine whether the observed differences in the distributions in Figures 3.2-3.5 were statistically significant, Pearson's chi-squared tests were conducted. The null hypothesis (H_0) was that the observed distribution as output by the model was drawn from the expected distribution:

H_0 : The distribution of deceased/injured Sims as estimated by the Markov model was drawn from the expected distribution;

H_1 : The distribution of deceased/injured Sims as estimated by the Markov model was NOT drawn from the expected distribution.

The test statistics and p -values for the chi-squared tests that were conducted for each of the three cases are summarised in Table 3.2. The high p -values indicated that the

Case	KS test statistic	p -value	H_0 rejected
Tumour absent, zero dose	21.9	0.99	No
Tumour present, zero dose	5.2	1.0	No
Tumour absent, uniform dose	49.0	0.55	No

TABLE 3.2: Summary of the Chi-squared test statistics. For these tests, the number of degrees of freedom was set at 51 as there were 52 bins in total (the example patient was initially aged 48, with 52 years remaining until they reached 100).

observed and expected distributions agreed in each case, and that the model was behaving as expected.

3.5 Discussion

The toolkit has been presented and output demonstrated in the publication. Verification of the model has also been carried out, with the result being that the model performed as intended. However, it must be stressed that this method is distinct from validation whereby the results are compared with outcomes of real patients. This step is vital before the model is able to be implemented clinically.

The technical aspects that were not presented in the publication also have limitations. Monte Carlo methods are inherently slow as a large number of iterations are required for accurate results. Although this process may be easily parallelised using packages such as OpenMP (OpenMP Architecture Review Board; www.openmp.org), the efficiency will be limited by the number of available processors. The model is currently only a prototype, but the intention is that it will be implemented clinically in the future. It is likely that many clinics will be publicly funded and may not have access to high performance computing resources that offer a larger number of processors.

The uncertainty associated with the model output is statistical uncertainty due to the Monte Carlo simulation being based on random sampling. There is no consideration of uncertainties in the model parameters or in the dose delivered to the tumour and organs at risk. These uncertainties have the potential to influence the predicted benefit (or lack thereof) of treating a patient with proton therapy.

Another limitation of this model is the assumption that the death probabilities (as a result of cancers and injuries) are dependent on time. However, it is more likely that the probability of death would depend on stage of cancer or grade of injury. The quality of life associated with an Unwell state would also depend on the grade or stage.

While the importance of validation has been addressed in the Discussion of the publication P1, it would likely be difficult to validate some aspects of the model. For example, it is difficult to know exactly what causes a second cancer as it is a biologically stochastic process.

3.6 Conclusion

The initial model has been developed successfully to consistently predict a clinical outcome given a radiotherapy treatment plan using a Monte Carlo approach. Although the model has incorporated numerous technical details, there are limitations that are addressed in a later model (see Chapter 4). However, if computational efficiency were not an issue, the presented model could potentially be implemented clinically upon external validation.

Chapter 4

Analytic Evaluation

4.1 Introduction

The initial model developed as part of this work was presented in Chapter 3. This model, referred to in this chapter as the Monte Carlo Evaluated (MCE) model, had limitations which were discussed. A notable limitation was the large computation time. In addition, there was no consideration of uncertainties associated with the input parameters or dose data used to calculate the state transition probabilities. It was concluded that the most efficient way to include these uncertainties was to evaluate the Markov model analytically instead of with a Monte Carlo simulation. Hence, the model presented in this chapter was developed to facilitate the inclusion of non-statistical uncertainties in the model estimates. These uncertainties, which can potentially impact on whether a patient is selected for proton therapy, are considered in a later version (see Chapter 5).

In this chapter, an updated model is presented and evaluated analytically, producing results consistent with the results of the model outlined in Chapter 3 (provided a sufficiently large number of Monte Carlo iterations). This model, referred to in this chapter as the Analytically Evaluated (AE) model, involves considering simultaneously the entirety of a hypothetical cohort of Sims (each being an identical copy of an individual patient under consideration), rather than each Sim in the cohort individually with each being a Monte Carlo simulation. For this reason, it is possible to obtain results with significantly

reduced computation time compared with the MCE model. However, a disadvantage is a greater memory usage requirement. At any given time, only the distribution of Sims among the Markov states is known, and it is not possible to track individual Sims in the AE model. Therefore, alternative methods must be employed to obtain results that are equivalent to the MCE model and to ensure efficient usage of memory. Both models were written in the C programming language.

The development of the AE model involved implementing several alterations in the Monte Carlo code, predominantly to increase computation efficiency and to provide a more realistic model output. Details of these are outlined in Section 4.2, where key differences in the Markov states and transition probabilities are described. Next, the method of evaluating the AE model is presented in Sections 4.3 and 4.4. The output is described in Section 4.5 and the results and computational efficiency are compared with the MCE model in Section 4.6. Finally, a discussion of the strengths and limitations of the AE model is provided in Section 4.7, with several points also being applicable to the MCE model.

4.2 Description of the AE model

While the two models are largely equivalent, there are key differences which are summarised in this section. More realistic predictions are facilitated in the AE model by incorporating the effects of differing injury grades and cancer stages on the quality of life of the patient. This allows for a more accurate estimate of the quality of life of a patient. Grades specify the severity of an injury, with higher grades corresponding to more severe injuries, which have a larger impact on quality of life. Different grades may also have different death and recovery probabilities (see Section 4.2.3). The staging of cancers is explained in Section 4.2.1. For each cancer, a state was allocated to each possible stage, and for each injury, a state was allocated to each possible grade. As a result, the number of Markov states in the AE model is significantly greater than that of the MCE model (see Section 4.2.6). The increased computational efficiency that could be achieved with

an analytic approach compared with a Monte Carlo simulation allowed for the inclusion of these additional states.

However, while the analytic evaluation is computationally more efficient compared with the Monte Carlo simulation (see Section 4.4), the computation speed is still a function of the number of Markov states. To further increase computational efficiency, techniques were employed to remove unnecessary Markov states. These included the removal of negligible probabilities, removing negligible injury grades, and combining all second cancer states into a single second cancer state.

4.2.1 Cancer staging

Cancers can be classified based on the likely prognosis [85]. Factors including tumour extent, lymph node involvement and the extent of metastasis affect the classification. The cancer classification groups are known as stages which are described as follows:

- Stage 0: Carcinoma in situ with no metastatic potential,
- Stage I: Smaller and less invasive cancers with negative nodes,
- Stages II and III: Increasing tumour or nodal extent,
- Stage IV: Distant metastasis.

A state was allocated to each possible stage of cancer. It was assumed that there were three stages in the model, as Stage 0 was not considered to impact quality of life. However, this assumption can be easily modified if necessary.

As prognosis is represented by the stage, it was concluded that death probability should depend on the stage for a given cancer, rather than the time since treatment. Therefore, each stage has a unique death probability. This is discussed in more detail in Section 4.2.3.

4.2.2 Inclusion of higher injury grades

Unlike the transitions from the Well state to states with higher cancer stages, it is possible to transition to states with higher injury grades directly from the Well state. The probability of this transition is given by the NTCP for each grade of injury. The events of developing different grades of the same injury are mutually exclusive, that is, it is not possible to be simultaneously affected by two grades of the same injury at the same time. Therefore, there are no Markov states where two grades of the same injury co-exist. It is also impossible to develop a grade of an injury once another has been recovered from (see Section 4.3.1). Note that there is a single Deceased state for each injury: death due to differing grades of the same injury will result in a transition to the same Deceased state.

If a member of the cohort does not have a given injury ℓ at a given time, then the probability of not developing that injury, $Pr^c(\ell, x)$, is the complement of the probability of developing any grade of that injury at year x , given by

$$Pr^c(\ell, x) = 1 - \sum_{i=1}^{N_g} NTCP_i(x), \quad (4.1)$$

where $NTCP_i(x)$ denotes the probability of developing grade i of the injury, and N_g denotes the total number of grades of the injury. This probability is implemented when calculating the transition probabilities with (3.5) and (3.6), which combine the probabilities of various events that constitute a transition. For example, consider the case where there are two possible injuries that could affect a patient, with Injury 1 having one grade and Injury 2 having two grades, a and b . For simplicity, death is not possible in this example. The probability of staying in the Well state (with row $i = 1$ and column $j = 1$) at year x is the probability of not developing any grades of any injury,

$$p_{11}(x) = (1 - NTCP_1(x))(1 - NTCP_{2a}(x) - NTCP_{2b}(x)). \quad (4.2)$$

It is also assumed that any time spent in lower grades of the same injury while recovering from a higher grade is negligible. Therefore, it is not possible to transition between different grades of the same injury.

If a constant number of grades is assumed for all injuries, then the result is a larger number of Markov states, corresponding to an increased computation time. As an alternative, the number of grades for each injury was defined as a variable that depends on the injury, as specified by the user. For example, Injury 1 may have a single grade while Injury 2 may have two.

Once the injury grades were incorporated into the model, the output was tested under a variety of conditions to ensure that the model was behaving as expected. These included varying the number of injuries and the number of grades for each. Death and injury recovery were also included. Example results are provided in Figures 4.1 and 4.2. It is evident that the model is behaving as expected after the inclusion of injury grades.

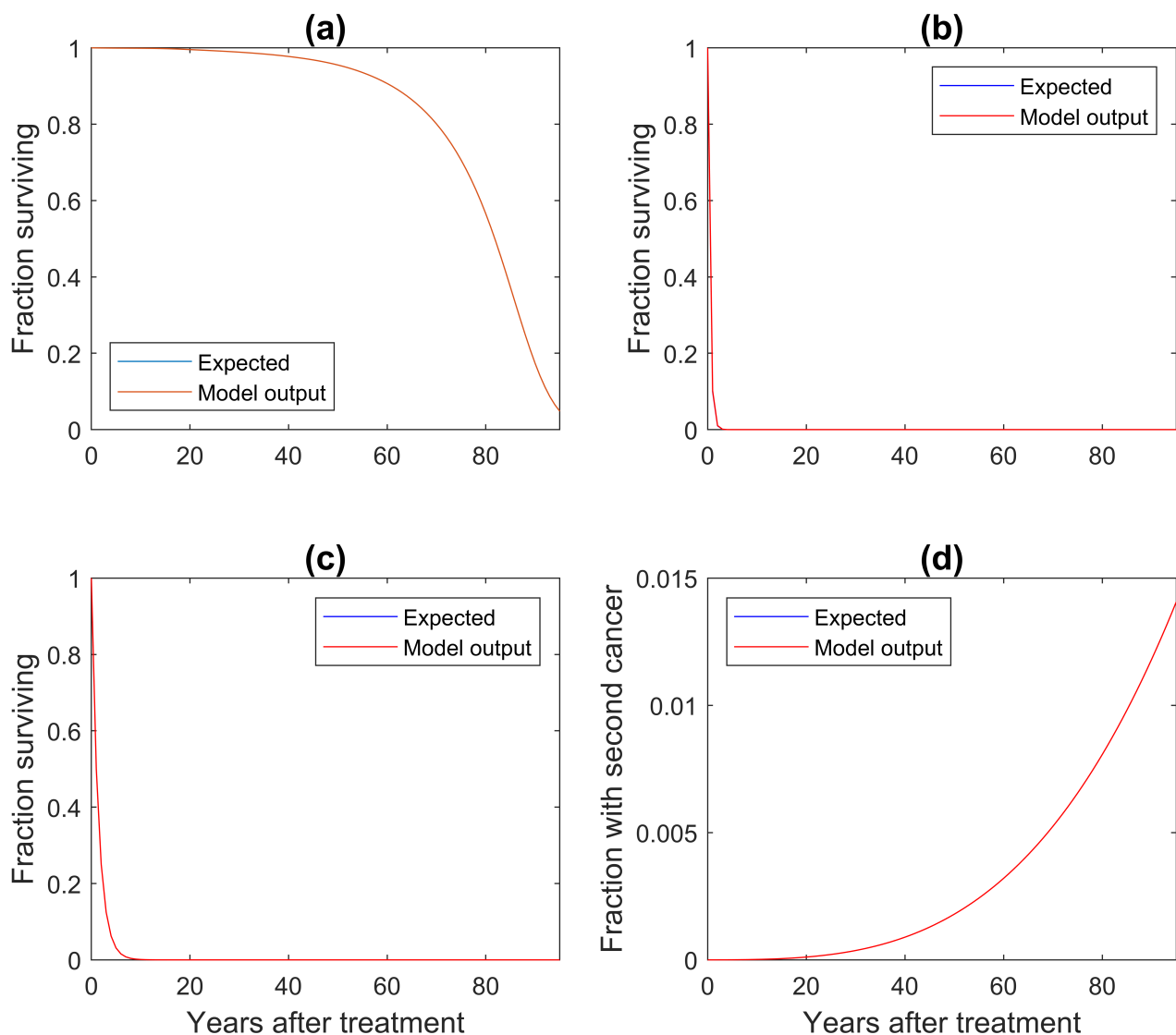


FIGURE 4.1: A comparison of the model output with expectations for each year after treatment and a variety of scenarios: (a) the surviving fraction with background death being the only possible cause of death (with a non-zero SPCIP and NTCP and a varying number of injuries and grades, which allows the transitions between these states to be checked); (b) the same as (a) with the only cause of death being the initial primary cancer (with probability 0.9 and a TCP = 1); (c) the same as (a) with the only cause of death being second cancer (with probability 0.5); (d) the fraction with a second cancer. In (d) all possible causes of death were disabled to allow this specific component of the model to be tested. A perfect agreement with expectations is apparent.

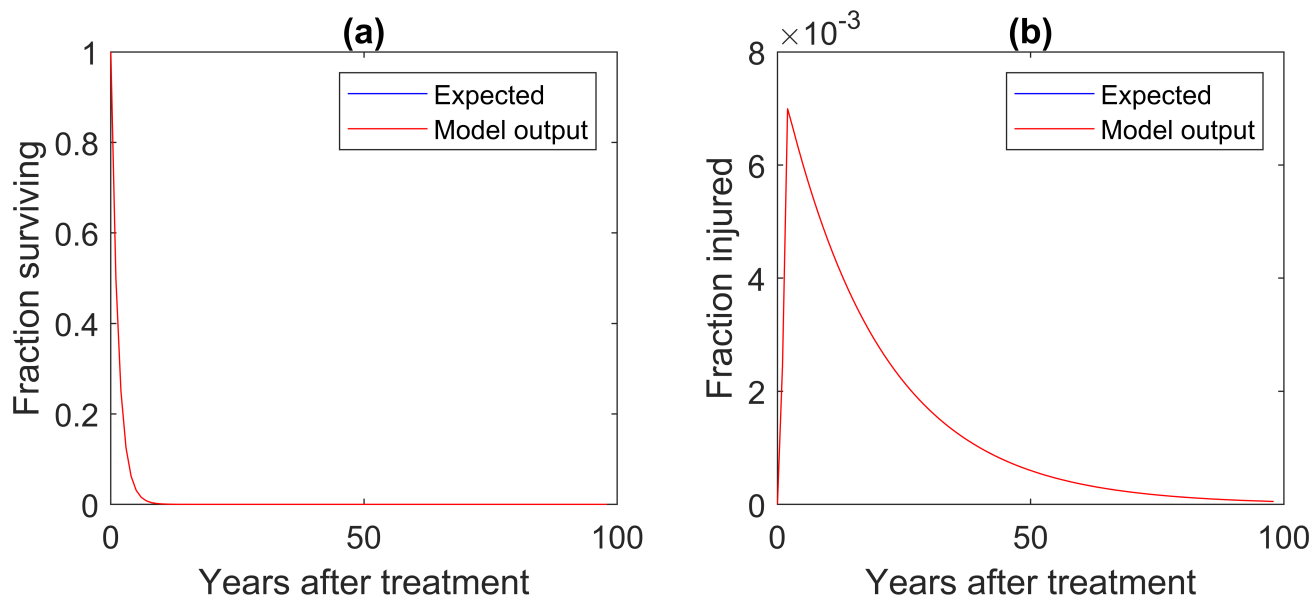


FIGURE 4.2: A comparison of the model output with expectations for each year after treatment and two injury related scenarios: (a) The fraction surviving with the only cause of death being injury related (with probability 0.5); (b) The fraction of the cohort with an injury where death is not possible and recovery occurs with a yearly probability of 0.05. Several tests were carried out for each scenario with varying grades of injuries.

A perfect agreement with expectations is apparent.

4.2.3 Grade/stage dependence of death probabilities

A key difference between the two models is that the recovery and death probabilities were assumed to be time-independent in the AE model, with the exception being the probability of unrelated death. This is time-dependent in both models, as it is not realistic to assume a constant unrelated background death probability for all ages.

Instead of being time-dependent, the cancer and injury death probabilities are grade- or stage-dependent in the AE model. This is a realistic assumption, as explained in Section 4.2.1. For the case of cancer, it could be argued that the cancer stage and the duration for which it has been present are correlated, but it was concluded that the stage of a cancer is a more appropriate determinant of death probability as the prognosis is represented by the stage.

Besides injury grades and cancer stages being more realistic determinants of the patient's treatment outcome than time, they also allow for a much smaller number of Markov

states. Recall, from Section 3.3.2.3, that the memoryless property of Markov models would mean that each state would be required to include the length of time a given injury or cancer has been present to allow the correct time-dependent probability to be applied. This would lead to an impractically large number of Markov states.

The death probabilities can be specified by the user. It is also possible to transition to higher stages of cancer from lower stages, with these probabilities also being user-defined. It is assumed that if a cancer develops as specified by the SPCIP, then the lowest possible stage of cancer develops and that the patient moves to the corresponding state. Refer to Section 4.2.2 for an explanation of the relationship between states representing different grades of injury.

4.2.4 Removal of negligible probabilities

The computation time and memory usage both increase rapidly with the number of OARs that are considered in the model. To increase efficiency, OARs that received a particularly small dose of radiation were not considered. For simplicity, this section will focus on the NTCP but similar methods were applied to the SPCIPs.

The all-time NTCPs are firstly calculated for each OAR. If the NTCP is less than a user-defined threshold, then the OAR is removed from the list of injuries. The removal, if required, is carried out before the time-dependent NTCPs are calculated using the method described in Section 2.4.2 of the publication P1. Ideally, the optimum threshold for probability exclusion would be determined through a sensitivity analysis to ensure that the accuracy of the output is not compromised by the removal of organs with small NTCPs.

4.2.5 Combining individual organ SPCIPs

Although it is important to consider the effect of a radiation-induced second cancer in the model, injuries typically have a greater effect on the quality of life of a patient as SPCs typically occur much later than injuries. It is therefore reasonable to assume that all

SPCs have an equal impact on the quality of life of a patient. As a result, a simplification was made by having only one Markov state to represent all of the SPCs that could potentially develop within the patient, rather than a separate state for each organ. This simplification results in a greatly reduced number of Markov states and a corresponding decrease in computation time. Note that a state still exists for the cases where a SPC is present in addition to every possible combination of injuries as well as the initial primary cancer.

The probability of transitioning to the combined second cancer state is given by the combined yearly SPCIP for each cancer site under consideration. The formula used to combine the probabilities is given by

$$\begin{aligned} Pr(E_1 \cup E_2 \cup \dots \cup E_N) &= 1 - Pr(E_1^c \cap E_2^c \cap \dots \cap E_N^c) \\ &= 1 - \prod_{i=1}^N [1 - Pr(E_i)], \end{aligned} \tag{4.3}$$

which gives the probability of developing a second cancer in tissue i (event E_i) at any year after treatment. Here, $Pr(E_i)$ is the SPCIP for tissue i for a particular year, N is the total number of tissues being considered as potential cancer sites, and E_i^c denotes the complement of E_i . It was assumed that the SPCIPs for each tissue were independent. This calculation is repeated for each year after treatment.

A single quality of life utility is assigned to this state and there is also a fixed (user-defined) probability of transitioning to higher grades of the cancer. This allows a reduced quality of life to be applied to a higher grade and for an increased death probability. Recovery from cancer remains impossible in the model. The exception is that it is possible to recover from the primary cancer immediately after treatment and before the beginning of the first Markov cycle.

4.2.6 Number of Markov states

There is a state for every possible combination of injury grade and cancer stage. The total number of states, without any minimisation techniques, is given by

$$n = (u + 2)^M (v + 1)^{N+1} + M + N + 2, \quad (4.4)$$

where M is the total number of injuries, u is the number of grades of each injury, N is the number of second cancer sites being considered and v is the number of cancer stages. The first summand on the right hand side of (4.4) represents the number of living states which is equivalent to the number of permutations, where the order matters. Each cancer and injury can take a particular ‘value’, depending on whether it is present in a given state, and which grade or stage if present. The combination of all these values in a given state corresponds to a particular permutation. The order matters as it is important that the grade/stage of a particular injury or cancer is known (rather than simply knowing that a certain number of injuries were of a given grade).

For a given state, each injury can have $u + 2$ possible ‘values’, one for each grade, one for the case where the injury is not present, and one for the case where the injury is not present but there is a history of the injury in the time since radiation treatment. This is explained in more detail in Section 4.3.1. Hence there are $u + 2$ values to select from, and M are selected. Similar reasoning applies to the cancers, however it is not possible to recover from cancers and hence there are only $v + 1$ values, and $N + 1$ are selected (N second cancers plus the initial primary cancer).

The addition of $M + N + 2$ represents the dead states, one for death from each second cancer and injury, one from the initial primary cancer and one from background death.

If the number of injury grades is variable for each injury (Section 4.2.2) and all possible second cancer sites are combined into a single site (Section 4.2.5), then the number of Markov states becomes

$$n = (v + 1)^2 \prod_{\ell=1}^M (u_{\ell} + 2) + M + 3, \quad (4.5)$$

where u_ℓ is the number of injury grades for injury ℓ . Here, N has been replaced with the value 1.

4.3 Evaluation

For the AE model, the transition matrix \mathbf{P} for a given year after treatment is constructed which stores the transition probabilities for each state pair at a particular year. A difference to the MCE model is that the entirety of the transition matrix must be evaluated at a given year, rather than a single row.

The row vector $\boldsymbol{\pi}_i$ represents the fraction of the cohort in each state at the end of year i and has n elements, one for each state. The first two elements represent the fraction of the cohort that are well and the fraction with the initial primary cancer only, respectively (see (4.8)). All other elements are zero at this point. At year i ,

$$\boldsymbol{\pi}_i = \boldsymbol{\pi}_{i-1} \mathbf{P}_i. \quad (4.6)$$

Thus, $\boldsymbol{\pi}_x$ at year x after treatment can be calculated by

$$\boldsymbol{\pi}_x = \boldsymbol{\pi}_0 \prod_{i=1}^x \mathbf{P}_i, \quad (4.7)$$

where $\boldsymbol{\pi}_0$ is given by

$$\boldsymbol{\pi}_0 = \left[TCP \quad (1 - TCP) \quad 0 \quad \dots \quad 0 \right]. \quad (4.8)$$

4.3.1 Preventing redevelopment of injuries

Including the possibility of recovery from an injury allows for a more accurate estimate of the quality of life of a patient. Biologically, it is impossible to develop a radiation-induced injury after it has been recovered from. It is therefore important to distinguish between

situations where a patient is well without a history of the injury (since treatment), and where he/she is well but has developed and recovered from an injury since the treatment.

In the MCE model, flags were used to determine which injuries had been recovered from and therefore which NTCPs should be zero in subsequent years (see Section 3.3.2.3). This is not possible in the AE model, however, as it is not possible to track individual Sims. In the AE model, an extra state was created to represent the situation where an injury had been recovered from to track the injury history. Transition probabilities from recovered states to injured states (associated with the same injury) are zero. For the case where multiple injuries could possibly affect the patient, an extra Markov state was created for every possible combination of injuries that could have been recovered from. An example of this is provided in the publication P2, which forms the basis of Chapter 5.

The possibility of redevelopment of cancers is not considered, as it is assumed that it is not possible to recover from cancers in this model.

4.3.2 Representing states

As with the MCE model, efficiently constructing the transition matrix for the AE model is a complex task. The reason for this complexity is that not only is there a state for every possible combination of cancer and injury that could be present, but also every possible combination of injury grades and cancer stages, which corresponds to a much larger number of states.

A similar method to the MCE model was employed to address this issue. A vector \mathbf{S} was used to represent each Markov state, with each element being an integer value. The combination of element values is unique to each state (see Section 3.3.3.1). There is an element to represent the presence or absence of each injury or cancer in a given state. The format of \mathbf{S} can be explained as follows: The first M elements represent the injuries, the second to last element represents the SPC and the final element represents the initial primary cancer. For example, if $M = 3$ and the initial primary cancer was treated successfully, then the state where the patient has grade 1 of injury 1, does not have injury 2, has grade 2 of injury 3, and has stage 2 SPC, is represented by $\mathbf{S} = [1 \ 0 \ 2 \ 2 \ 0]$.

In the AE model, the form of \mathbf{S} has two key differences compared with the form in the MCE model:

1. The number of elements.
 - (a) All second cancers are combined into a single state (see Section 4.2.5), and hence there is only one element in \mathbf{S} to represent the presence or absence of second cancers.
 - (b) An element is also added to represent the initial primary cancer. In the MCE model, each Sim is tracked separately and hence it is possible to simply use a flag to mark the presence or absence of a successful treatment. However, this is not possible in the AE model as the cohort is considered simultaneously, and it is therefore necessary to track the status of the initial primary cancer using other means. Consequently, there are $M + 2$ elements in each vector \mathbf{S} . For example, if $M = 3$ and $N = 2$, in the MCE model

$$\mathbf{S} = [s_1 \quad s_2 \quad s_3 \quad s_4 \quad s_5], \quad (4.9)$$

and in the AE model

$$\mathbf{S} = [s_1 \quad s_2 \quad s_3 \quad s_4 \quad s_5], \quad (4.10)$$

where the positions of the elements representing second primary cancer states are highlighted in pink and the initial primary cancer in yellow.

2. The possible values for the elements. In the MCE model, if an injury or cancer is present, then the value of its corresponding element is 1. Otherwise, it is 0. In the AE model the state at the beginning of a cycle is represented by

$$\mathbf{S}_{i,\ell} = \begin{cases} r & \text{if patient has grade/stage } r \text{ of injury/SPC } \ell, \\ -1 & \text{if patient has a history of injury } \ell, \\ 0 & \text{otherwise.} \end{cases} \quad (4.11)$$

The state at the end of the cycle is represented by \mathbf{S}_j , which has each element ℓ defined in the same manner:

- (a) Integers greater than 1 are used to account for states where there are higher grades of injury or stages of cancer present. For example, if the lowest grade/stage is present, then the element representing the corresponding injury/cancer is assigned a value of 1. If grade/stage 2 is present, it is assigned a value of 2 and so on.
- (b) Recovered states. Additional states, called recovered states, were assigned to represent situations where there has been a history of a particular injury (see Section 4.3.1). States where an injury is absent but has been present in the past have the element representing the injury assigned to -1 .

4.3.3 Structure of the transition matrix

The row order of the transition matrix for each year is as follows:

1. The Well state is listed first, followed by the initial primary cancer only (no second cancers or injuries) state.
2. All Unwell states are listed with the exception of the states where an injury has been recovered from. These states are listed in lexicographical order. For example, the order of the first states (represented by \mathbf{S}) are: $[0\ 0\ 0\ 0]$ (well), $[0\ 0\ 0\ 1]$, $[0\ 0\ 0\ 2]$, $[0\ 0\ 0\ 3]$, $[0\ 0\ 1\ 0]$, etc.
3. These Unwell states are followed by the Unwell states that are duplicated for the cases where there has been a history for one or more injuries. For a given number of absent injuries, there is a duplicate for every possible combination of histories of the absent injuries. For example, if there are two absent injuries, Injuries 1 and 2, then there are four states with only these injuries absent (all other factors constant, including the presence or absence of cancers and other injuries). If the cancers are absent, then there is the state for where both have never been present, represented by $\mathbf{S} = [0\ 0\ 0\ 0]$, a state for where both have been present ($\mathbf{S} =$

$[-1 \ -1 \ 0 \ 0]$) and two states for the cases where one has been present in the past ($\mathbf{S} = [0 \ -1 \ 0 \ 0]$ and $\mathbf{S} = [-1 \ 0 \ 0 \ 0]$). Of these four, the first is listed with the non-recovered Unwell states in the previous item.

4. The order of the recovered Unwell states follows the order of the non-recovered Unwell states, that is, all duplicates of a given non-recovered state are listed together. The order of these follows a binary pattern, with a 1 being replaced with a -1 to denote the history of an injury. For example, consider the simple case where there are no cancers present and two injuries. The order of the states are demonstrated as follows for the case where the cancers are absent, along with a description of each of the states:

$[0 \ 0 \ 0 \ 0] \rightarrow$ Well, no injury history,
 $[0 \ 1 \ 0 \ 0] \rightarrow$ Injury 2, no history of Injury 1,
 $[1 \ 0 \ 0 \ 0] \rightarrow$ Injury 1, no history of Injury 2,
 $[1 \ 1 \ 0 \ 0] \rightarrow$ Injuries 1 and 2,
 $[0 \ -1 \ 0 \ 0] \rightarrow$ Well, history of Injury 2,
 $[-1 \ 0 \ 0 \ 0] \rightarrow$ Well, history of Injury 1,
 $[-1 \ -1 \ 0 \ 0] \rightarrow$ Well, history of both injuries,
 $[1 \ -1 \ 0 \ 0] \rightarrow$ Injury 1, history of injury 2,
 $[-1 \ 1 \ 0 \ 0] \rightarrow$ Injury 2, history of injury 1.

5. Finally, the Deceased states are listed, one for each injury, followed by the second cancer Deceased state, the initial primary cancer Deceased state and the background Deceased state.

4.3.4 Calculation of transition probabilities

The elements of the transition matrix are calculated in the AE model using a similar method to the MCE model (see Section 3.3.3.2). Development of different injuries and

SPCs are independent (with the exception of differing grades of the same injury). Therefore, the probability of transitioning from state i to state j at year x after treatment is given by

$$p_{ij}(x) = (1 - Pr(\text{Die}, i, x)) \prod_{\ell=1}^{M+2} Pr(E_{\ell}, i, j, x), \quad i = 1, \dots, n - n_d, \quad j = 1, \dots, n - n_d, \quad (4.12)$$

where $Pr(E_{\ell})$ is the probability of event E_{ℓ} occurring and $n_d = M - 3$ is the number of Deceased states. The events associated with a given transition are determined by comparing the state at the beginning of the cycle, represented by \mathbf{S}_i , to the state at the end of the cycle, represented by \mathbf{S}_j . The values of each of the ℓ elements of \mathbf{S}_i and \mathbf{S}_j are determined by the events that occur in the transition. The events include injury or second cancer development, or injury recovery, with probabilities

$$Pr(E_{\ell}, x) = \begin{cases} \text{NTCP}_{\ell,r}(x) & \text{if } \ell \leq M, \quad \mathbf{S}_{i,\ell} = 0, \quad \mathbf{S}_{j,\ell} = r, \quad r \geq 1, \\ \text{SPCIP}(x) & \text{if } \ell = M + 1, \quad \mathbf{S}_{i,\ell} = 0, \quad \mathbf{S}_{j,\ell} = 1, \\ 1 - \sum_{r=1}^u \text{NTCP}_{\ell,r}(x) & \text{if } \ell \leq M, \quad \mathbf{S}_{i,\ell} = 0, \quad \mathbf{S}_{j,\ell} = 0, \\ 1 - \text{SPCIP}(x) & \text{if } \ell = M + 1, \quad \mathbf{S}_{i,\ell} = 0, \quad \mathbf{S}_{j,\ell} = 0, \\ Pr(E_{IR,\ell,r}) & \text{if } \ell \leq M, \quad \mathbf{S}_{i,\ell} = r, \quad r \geq 1, \quad \mathbf{S}_{j,\ell} = 0, \\ 1 - Pr(E_{IR,\ell,r}) & \text{if } \ell \leq M, \quad \mathbf{S}_{i,\ell} = \mathbf{S}_{j,v} \geq 1, \\ Pr(E_{SI,s}, i_s, j_s) & \text{if } \ell = M + 1, \quad \mathbf{S}_{i,\ell} \geq 1, \quad i_s < j_s, \\ 1 - Pr(E_{SI,s}, i_s, j_s) & \text{if } \ell = M + 1, \quad \mathbf{S}_{i,\ell} = \mathbf{S}_{j,\ell} \geq 1, \quad i_s < j_s, \\ Pr(E_{SI,p}, i_s, j_s) & \text{if } \ell = M + 2, \quad \mathbf{S}_{i,\ell} \geq 1, \quad i_s < j_s, \\ 1 - Pr(E_{SI,p}, i_s, j_s) & \text{if } \ell = M + 2, \quad \mathbf{S}_{i,\ell} = \mathbf{S}_{j,\ell} \geq 1, \quad i_s < j_s, \end{cases} \quad (4.13)$$

where $Pr(E_{SI,s}, i_s, j_s)$ is the time-independent probability of the stage of second cancer increasing from stage i_s to stage j_s , $Pr(E_{SI,p}, i_s, j_s)$ is the time-independent probability of the stage of initial primary cancer increasing from stage i_s to stage j_s , $\text{NTCP}_{\ell,r}$ is the probability of developing grade r of injury ℓ , $Pr(E_{IR,\ell,r})$ is the probability of recovering from grade r of injury ℓ , and u is the total number of grades of injury ℓ . The probability

of the stage of cancer increasing can be unique for a given pair of stages. Note that the injury recovery is time-independent.

Many of the transitions from \mathbf{S}_i to \mathbf{S}_j are impossible, including:

- $\mathbf{S}_{i,\ell} = -1, \mathbf{S}_{j,\ell} > 0, \ell \leq M$: developing an injury again once recovered from,
- $\mathbf{S}_{i,\ell} \neq 0, \mathbf{S}_{j,\ell} = 0, \ell \leq M$: erasing a history of injury,
- $\mathbf{S}_{i,\ell} = 0, \mathbf{S}_{j,\ell} = -1, \ell \leq M$: recovering from an injury before obtaining it,
- $\mathbf{S}_{i,\ell} > 0, \mathbf{S}_{j,\ell} > 0, \mathbf{S}_{i,\ell} \neq \mathbf{S}_{j,\ell}, \ell \leq M$: transitioning between different grades of the same injury,
- $\mathbf{S}_{i,\ell} > 0, \mathbf{S}_{j,\ell} = 0, \ell > M + 1$: cancer recovery,
- $\mathbf{S}_{i,\ell} = 0, \mathbf{S}_{j,\ell} > 1, \ell = M + 1$: developing a higher stage of cancer before the lowest stage,
- $\mathbf{S}_{i,\ell} = 0, \mathbf{S}_{j,\ell} > 1, \ell = M + 2$: recurrence of the initial primary cancer.

4.4 Computational efficiency methods

With a state for every possible combination of injury grades, cancer stages and also the recovered states, a challenge exists in storing the entirety of the transition matrix simultaneously. As explained in Section 4.3.4, the transition matrix is a sparse matrix as many elements are zero. This property was utilised to allow for more efficient storage and computation time without compromising the model predictions. This is particularly important as the intended purpose of the model is to provide decision support in the clinic with minimal computation time. Furthermore, it is likely that many clinics would rely on standard desktop computers due to a limited access to supercomputer clusters.

4.4.1 Method of matrix condensation

Each matrix element is calculated in order of increasing row, then increasing column. The method of matrix storage involves storing only the non-zero values of the transition matrix as they are calculated. These values are stored sequentially in a single one-dimensional array in order of increasing row then increasing column (the order in which they are calculated). As a number of probabilities are time-dependent, the length of this array is not necessarily constant for each year after treatment. The array is accompanied by two other arrays of equal length: one stores the row index of each non-zero element and the other stores the column index of each non-zero element. The indices are important as they allow successive transition matrices to be multiplied together.

4.4.2 Filling of the transition matrix in blocks

The slowest component of the model evaluation is the process of determining which events contribute to each probability in the transition matrix (see Section 4.3.4). It is not possible to know which transitions are impossible until the events that contribute to the transition have been assessed. As the matrix is sparse, there are many elements that are calculated but not stored as they evaluate to zero. This represents unnecessary computation time.

Fortunately, the structure of the matrix has a pattern due to the lexicographical ordering of the states (see Section 4.3.3). This property was exploited to decrease the computation time required to construct the transition matrix for each year. For example, consider the case where there are three injuries and no cancers. The state where the first injury is present and the others are not is represented by $\mathbf{S}_i = [1\ 0\ 0\ 0\ 0]$. Suppose the elements of the row corresponding to this state are being calculated. The first transition considered is from this state to the Well state, represented by $\mathbf{S}_j = [0\ 0\ 0\ 0\ 0]$. This transition is impossible as recovery from the first injury would mean a transition to the state represented by $[-1\ 0\ 0\ 0\ 0]$. Hence, once the first element of \mathbf{S}_i in this case is compared with the first element of \mathbf{S}_j , the transition matrix element will evaluate to zero regardless of any events associated with the other injuries present. Therefore, it is not necessary to consider these

events. Similarly, it is impossible to transition to all other states where $\mathbf{S}_{j,1} = 0$ if $\mathbf{S}_{i,1} = 1$ and these probabilities are not stored as a result. Recall that, due to the lexicographical order, all states where $\mathbf{S}_{j,1} = 0$ (or where $\mathbf{S}_{j,1} = 1$) are listed consecutively due to the order of the states. States where $\mathbf{S}_{j,2} = 0$ or where $\mathbf{S}_{j,2} = 1$ are not necessarily listed together but are listed in sub-blocks. The sizes of the blocks decrease with increasing index value of the element of \mathbf{S}_j . Hence, the largest computation savings correspond to elements of \mathbf{S}_j with the smaller indices.

In general, the process of filling the matrix in blocks is as follows:

1. Each row is considered separately (i is kept constant), that is, the probabilities of transitioning from a given state to all other states. The rows representing Deceased states are not included in the algorithm as all elements of these rows will be zeros (except the diagonal elements, $i = j > n - n_d$, which are equal to 1 as it is impossible to leave a Deceased state). Once rows $i \leq n - n_d$ have been filled, n_d elements with the value 1 are appended to the end of the array that contains the values of the transition matrix (see Section 4.4.1).
2. The events corresponding to the first transition probability in the row are determined by considering elements of $\mathbf{S}_{i,\ell}$ and $\mathbf{S}_{j,\ell}$ for $\ell = 1, \dots, M + 2$.
3. If a forbidden event is encountered for a given i and j , then the transition probability is assigned a value of 0 regardless of whether all possible values of ℓ have been considered. This is because each transition consists of several events and if at least one event is forbidden, then the transition is not possible. Thus, it is not necessary to check all possible values of ℓ , thereby saving computation time.
4. A check is then carried out to determine whether the same forbidden event E_ℓ exists for a transition to be made from state i to the next state in the row, $j + 1$ (before checking if other events exist by keeping ℓ constant). If it does occur, then the transition is also forbidden and the state $j + 1$ is skipped without being stored. No other calculations are necessary to determine that transition probability. The reduced number of calculations results in a reduced computation time.

5. Each of the following states are skipped ($j + 1$ is incremented while keeping ℓ constant) until a state is encountered where the forbidden event does not occur. Then Step 2 is repeated for the next probability in the row i .
6. The process is repeated until $j = n - n_d + 1$ and the absorbing states are reached. These are appended to the list of transition matrix values. If the row represents a state where an injury or cancer is not present, then the corresponding element is set to zero and not stored in the condensed representation of the transition matrix (recall that only the non-zero elements are stored).
7. The process is repeated for the next row of the transition matrix.

4.5 Model output

Evaluating the AE model analytically involves making calculations based on the distribution of the cohort among Markov states at each cycle. From here, it is possible to calculate both the raw and quality-adjusted life expectancy.

4.5.1 Raw life expectancy

Both mean and median raw life expectancy can be calculated from the fraction of the cohort that moved to a Deceased state each year. This fraction for a given year is the difference between the surviving fractions of two consecutive years. Initially, at $\tau = 0$, the surviving fraction is 1 by default as it is not possible to die at this time in the model. In a discrete-time Markov model, the time to death, τ , is a discrete random variable. Hence it is only possible for a member of the cohort to die at $\tau = 1, 2, 3, \dots$. If death occurs at $\tau = 1$ for example, then the life expectancy is 1.

The mean life expectancy is equivalent to the expected time until death, $E[\tau]$, plus the patient's age at treatment, a . The expected time to death is determined by finding the mean of the distribution of yearly deaths, which depends on the probability of dying in

a given year k , $Pr(\tau = k)$:

$$E[\tau] = \sum_{k=1}^m k Pr(\tau = k), \quad (4.14)$$

where $m = 100 - a$, $Pr(\tau = m) = 1 - \sum_{k=1}^{m-1} Pr(\tau = k)$ and $Pr(\tau > m) = 0$, that is, the surviving fraction of the cohort is absorbed after the final year of the Markov chain. If a patient does not die before $100 - a$, then it is assumed that they are deceased after $100 - a$, as it is not possible to live beyond the age of 100 in the model. This assumption allows comparison of the output with that of the MCE model, where each Sim is considered individually.

The probability of dying at a given year can be calculated as

$$Pr(\tau = k) = \boldsymbol{\pi}_0^A \prod_{i=0}^{k-1} \mathbf{P}_A^{(i)} \mathbf{d}, \quad (4.15)$$

where $\boldsymbol{\pi}_0^A$ contains the initial fraction of the cohort in each living state, the square matrix $\mathbf{P}_A^{(i)}$ is the upper left quadrant of the transition matrix at year i that represents the probabilities of transitioning between the living states, and \mathbf{d} contains each of the probabilities of moving from a living state to a Deceased state. The sum of the numbers of elements of \mathbf{d} and $\boldsymbol{\pi}_0^A$ is equal to the number of Markov states, n .

Thus, the mean life expectancy is calculated as

$$E[\tau] = \sum_{k=1}^m k \left(\boldsymbol{\pi}_0^A \prod_{i=1}^{k-1} \mathbf{P}_A^{(i)} \mathbf{d} \right). \quad (4.16)$$

It is also possible to calculate the mean “on the fly” with a less computationally expensive approach. The expectation time can be re-written as

$$E[\tau] = \sum_{k=1}^m k Pr(\tau = k)$$

$$\begin{aligned}
&= \sum_{k=1}^m \sum_{j=1}^k Pr(\tau = k) \\
&= \sum_{j=1}^m \sum_{k=j}^m Pr(\tau = k) \\
&= \sum_{j=1}^m Pr(\tau \geq j) \\
&= \sum_{j=0}^m Pr(\tau > j) \\
&= \sum_{j=0}^m \sum_{i=1}^n (\pi_j)_i,
\end{aligned} \tag{4.17}$$

where n is the number of Markov states. This is essentially the sum of the surviving fraction at each year.

In reality, the distribution of the time until death is a skewed distribution, as typically the death probability increases with time due to the effect of background death. Therefore, it is more appropriate to report the median life expectancy instead of the mean, as it generally is not distorted by skewed data. It is possible that two treatment plans could have equal mean life expectancies but unequal medians and therefore the median can provide additional information. By definition, half of the cohort will die before the median life expectancy. This quantity may therefore be calculated by finding the age corresponding to half of the area contained within the distribution of time until death.

4.5.2 Quality-adjusted life expectancy

The calculation of the QALE is more complicated than the raw life expectancy. Unlike the MCE model, it is not possible to track individual members of the cohort and it is therefore not straightforward to obtain a distribution of QALEs for the calculation of the mean and median QALE. For example, if a certain fraction moves to the Deceased state in a given year, each member in that fraction would not necessarily have had the same

injuries or cancers for the same amount of time and hence may not have had the same quality of life.

Equation (4.17) calculates the mean survival each year, and then sums over all years. The result is the mean life expectancy. A similar method can be used to obtain the mean QALE, by multiplying the fraction of the cohort in a given state by the utility of the corresponding state, and then summing over all years:

$$\widehat{\text{QALE}} = \sum_{j=0}^m \sum_{i=1}^n (\pi_j)_i \cdot u_i \quad (4.18)$$

where $\mathbf{U} = [u_1 \quad u_2 \quad \dots \quad u_s]$ is a vector of the quality of life utilities for each state. While the mean QALE is straightforward to obtain, this is not the case for the median QALE. Therefore, the median QALE was not considered in the AE model.

4.5.3 Other metrics

The proportion of the cohort that died of particular causes (discussed in Chapter 3) is evaluated by the AE model. This is achieved by having several Deceased states, one for each possible cause of death (as explained in Section 3.3.2.4), and considering the fraction of the cohort in each of these states when the time horizon is reached.

The yearly fraction of the surviving cohort that suffered an injury or second cancer after treatment is also evaluated by the model. This is calculated as the sum of the fractions of the cohort in each injured or cancer state, respectively (some states are both injured and cancer states). The fraction with an injury or second cancer is divided by the surviving fraction to give the surviving fraction in an unwell state.

4.6 Comparison with the MCE model

It was necessary to compare the output of the two models to ensure that the AE model was developed correctly. The output of the MCE model was verified to ensure that the

model was behaving as expected (see Section 3.4). Therefore, if the two model outputs agree, then it could be inferred that the AE model is also behaving as expected.

The fundamental difference between the evaluation of the two models is that the entirety of the hypothetical cohort is considered simultaneously in the AE model, whereas individual cohort members are considered separately in the MCE model. However, this should not affect the predictions of the model.

Another important difference between the evaluation of each model is the alternative methods used to prevent injury redevelopment. While this should not result in differing predictions, it does increase the number of states in the AE model compared with the Monte Carlo simulation. This is important to consider when comparing the computation time of the two models.

The death and recovery probabilities are time-independent in the AE model (with the exception of the background death probabilities), as described in Section 4.2.3. However, for the purposes of comparing the output between the two models, the death probabilities in the MCE model were manually set to time-independent values that were equal to those used in the AE model. Furthermore, for the purposes of comparing with the MCE output, injury grades and cancer stages were not included in the AE model as these were not originally considered in the MCE model.

The SPCIPs are combined for all organs in the AE model. Therefore, only a single organ was considered in the comparison to remove the effect of the alternative processing of these probabilities on the comparison. Negligible probabilities were not removed.

The data from a proton plan for a four-year-old female base of skull chordoma patient was considered for the test. This data was the input to both models. A cohort size of 10^6 was selected for all tests of the Monte Carlo simulation as this gave results consistent to 3 significant figures. The yearly primary cancer death probability was set to 0.9 and the yearly second cancer death probability was set to 0.1. The injury death probability was set to 0. The results are summarised in Table 4.1. The two models produce the same result to within 3 significant figures, and also agreed to within the statistical error of the MCE model. This is the same level of accuracy that can be obtained when comparing

different outputs of the Monte Carlo simulation calculated with the same input data. A comparison of the graphical output is given in Figures 4.3-4.5. These compare at each year after treatment the surviving fraction of the cohort, the fraction injured, and the fraction with a second cancer, respectively. Good agreement was achieved for all cases tested.

	Monte Carlo	Analytic
Mean raw LE (years)	67.43 ± 0.08	67.42
Mean QALE (years)	41.05 ± 0.32	41.00

TABLE 4.1: The results of each model given equivalent input data. The TCP resulting from the input treatment plan was 0.83 and the NCTP was 0.99. The Monte Carlo cohort size was 10^6 and the statistical uncertainty associated with the life expectancies is shown.

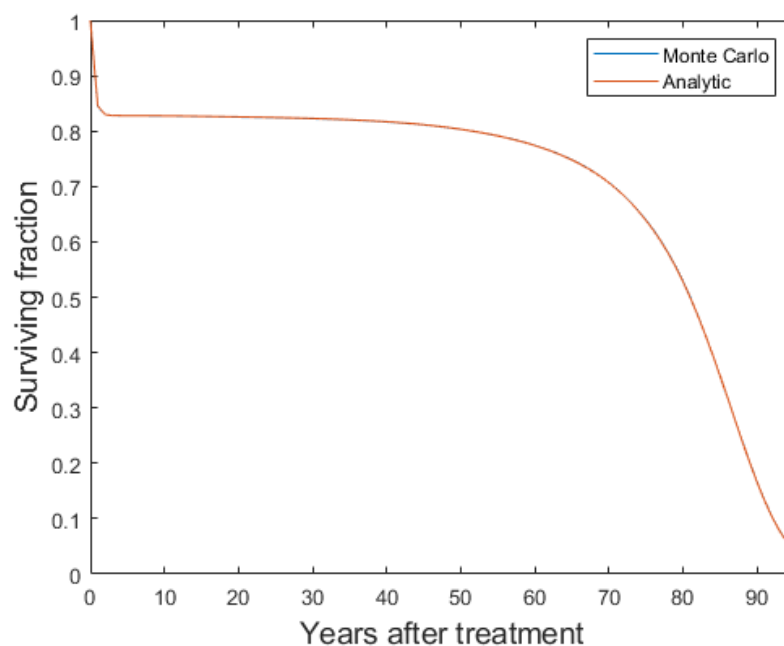


FIGURE 4.3: The fraction of the cohort surviving for each year after treatment. The cancer death probability was set at 0.9 which results in the initial sharp decrease in survival. The possibility of death due to injury was disabled. The background death accounts for the large decrease in survival in the later years after treatment. A good agreement between the two models is apparent.

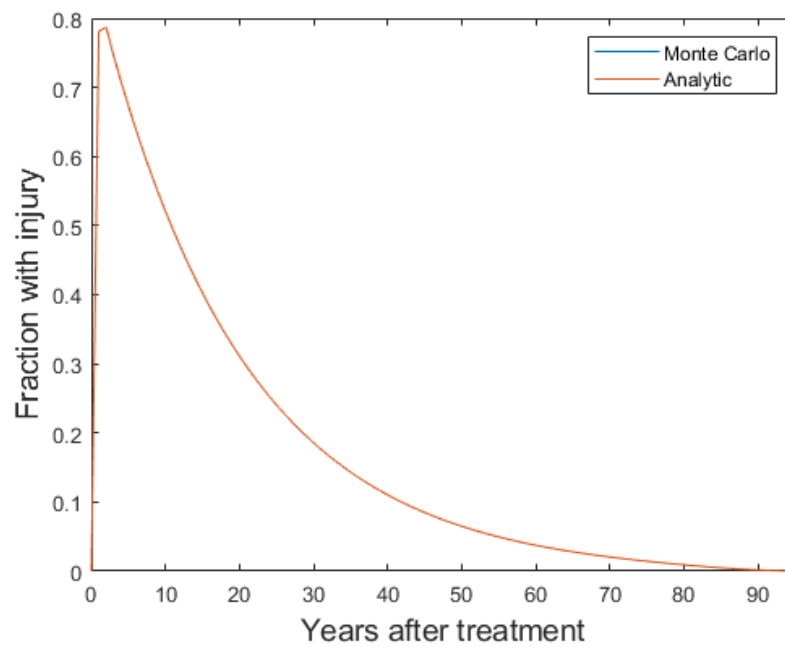


FIGURE 4.4: The fraction of the cohort with an injury for each year after treatment. The probability of injury recovery was set to 0.05. A good agreement between the two models is apparent.

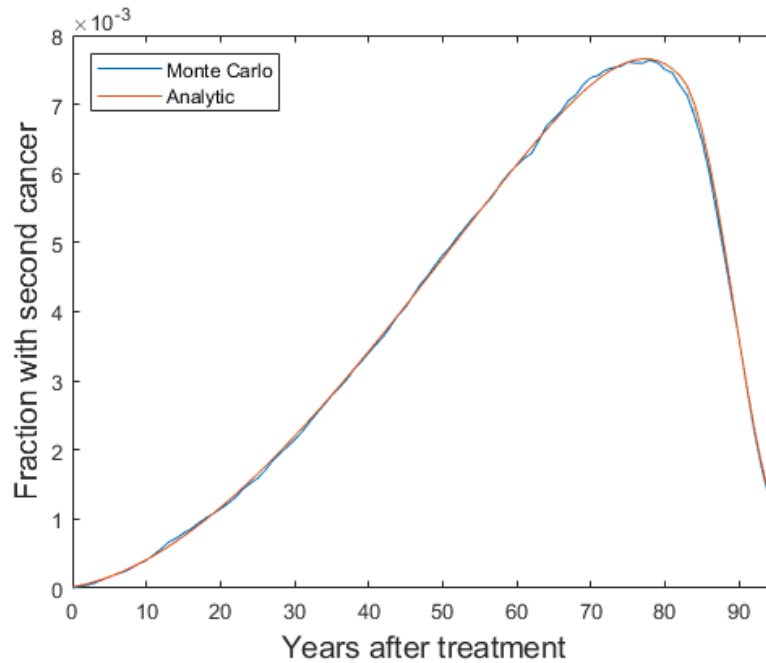


FIGURE 4.5: The fraction of the cohort with a second cancer at each year after treatment. The small discrepancy apparent is a result of the relatively small probability of developing a second cancer (differences are more easy to detect compared to Figure 4.4, due to the difference in vertical scale).

The computation time was also compared between the two models. For the Monte Carlo simulation, this time is a function of both the hypothetical cohort size (number of iterations) and the number of states. In contrast, the analytic computation time is function of the number of states only. The same patient data was used (proton treatment plan for the four-year-old female patient with base of skull chordoma). As stated above, the AE model combines all organs to give a single estimate of the SPCIP. Hence, only the length of the list of injuries was increased sequentially to determine the computational efficiency for an increasing number of states. It is unlikely that more than 6 OARs would receive a significant dose, so this was the maximum considered. The results are tabulated in Table 4.2, and graphed in Figure 4.6. Recall that additional states are required in the AE model (see Section 4.2.6).

# Injuries	MCE model		AE model	
	# States	Time (s)	# States	Time (s)
1	8	78.2	16	0.052
2	13	93.3	41	0.13
3	22	123.0	114	0.4
4	39	193.8	331	1.5
5	72	365.8	980	10.3
6	137	748.1	2925	125.1

TABLE 4.2: The results of the computation time comparison.

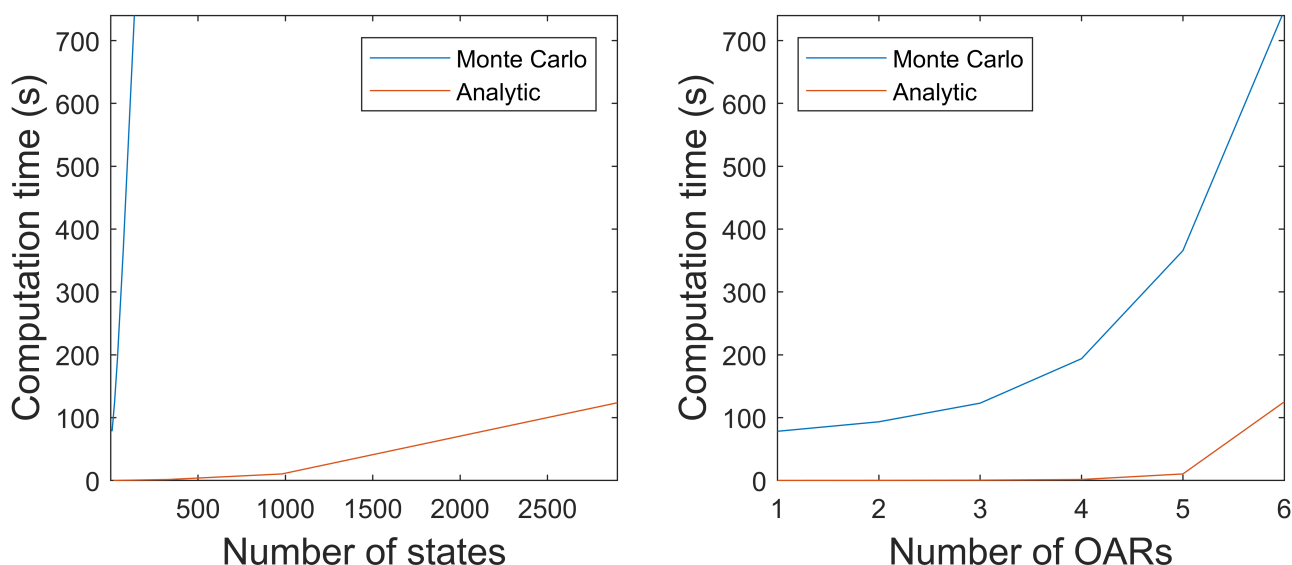


FIGURE 4.6: The computation time for each model, as a function of the number of Markov states (left) and organs at risk (right). For an equivalent number of states/organs at risk, the AE model clearly offers an advantage over the MCE model in terms of computation time.

The AE model offers a significantly reduced computation time compared with the Monte Carlo simulation. For six OARs, the runtime is six times faster even though the number of states is at least 20 times greater. When the number of states are approximately equal, the AE model is approximately 1,800 times faster.

It should be noted that it can be difficult to compare the two models fairly in terms of computation speed for several reasons. Firstly, the entire matrix is calculated each year in the AE model whereas only a single row is calculated each year in the MCE model. Secondly, a number of computational efficiency techniques were implemented in the AE model.

4.7 Discussion and conclusion

The AE model has been implemented successfully and produces predictions with a significantly reduced computation time compared with the MCE model. A major contributor to this is the efficiency of matrix multiplications compared with Monte Carlo simulations.

Furthermore, these results do not have the statistical uncertainty that is present in the results of the Monte Carlo simulation. However, like the MCE model, the AE model has limitations. Some of these are unique to the latter but many also apply to the former.

A key enhancement of the model is that the injury and cancer death probabilities depend on the grade or stage rather than time. While it is more realistic for death probabilities to depend on severity rather than time, a difficulty exists in sourcing appropriate death probabilities, particularly for different stages of cancer. Furthermore, there is a time- and dose-independent probability of transitioning to higher stages of cancer to allow for more realistic quality of life estimates for the fraction of the cohort being affected by each cancer. However, it can be difficult to source reliable estimates of these parameters. In the absence of reliable estimates, a single representative quality of life utility can be applied to the lowest stage while omitting the probabilities of transitioning to higher grades.

In other models developed for patient selection for proton therapy, patient-specific variables other than dose are included [66, 68, 69]. In this thesis (both the MCE and AE models), the only patient-specific variables are dose, age and gender. There is no consideration of concurrent chemotherapy, pre-radiotherapy patient history, or genetics, all of which may influence patient outcome after receiving radiotherapy. However, due to the structure of the model, it would be difficult to include these factors without having an unacceptably large number of Markov states.

In both models, there is uncertainty inherent in the application of the NTCP model parameters, as they have been determined in studies based on populations treated with photons only. As a result, they may not be able to produce reliable predictions for a patient when treated with proton therapy. In future work, the model parameters should be validated using populations treated with both photon therapy and proton therapy.

Finally, it is important to consider the quality of the input data. In both models, uncertainty associated with the input data is not taken into consideration. The methods used to include these uncertainties are outlined in Chapter 5.

Chapter 5

Model Uncertainties Incorporating Robust Plan Analysis

The publication P2 forms the basis of this chapter:

Austin, A.M., Douglass, M.J.J., Nguyen, G.T. & Penfold, SN. Patient selection for proton therapy: A radiobiological fuzzy Markov model incorporating robust plan analysis. *Mathematical Medicine and Biology*. 2019 (Submitted May 2019).

5.1 Introduction and motivation

The AE model presented in Chapter 4 was developed to improve computational efficiency and to include the effects of higher injury grades and cancer stages on patient quality of life. In addition, it facilitated the development of the model presented in this chapter, which forms the basis of the publication P2. This model incorporates the effects of uncertainties in the radiobiological model parameters and the dose delivery into its predictions.

Where different treatment modalities are being considered, there may be significant differences in uncertainties that could arise due to patient set up and changes to the treatment volume over the course of treatment. Internal organ motion/displacement and breathing

may exacerbate these effects. In addition, due to the prominence of the Bragg peak in the proton therapy dose deposition, the range uncertainty for proton therapy can be significantly larger than that of X-ray therapy. This is because a small shift in the Bragg peak can have a much more significant effect on the tumour and healthy tissue than changes in the X-ray therapy dose distribution. These effects are not taken into account in the AE model. It is important to consider these uncertainties as they can affect whether a patient is selected for proton therapy.

The aims of this publication were to present the model that incorporated these uncertainties into its predictions, and to demonstrate the output with an example patient. The model is based on a fuzzy Markov model, which is a Markov model with uncertain transition probabilities.

This chapter details features and technical aspects that were not included in the publication (Section 5.3), including the uncertainties in the quality of life utilities (Section 5.3.3). A discussion and conclusion are given in Section 5.4.

5.2 Statement of contribution

5.2.1 Conception

The idea to incorporate the effect of uncertainties into the model was first conceptualised by Scott Penfold. All authors contributed to the development of ideas and methods.

5.2.2 Realisation

The writing of the code and analysis was performed by Annabelle Austin, with advice provided by Scott Penfold, Michael Douglass and Giang Nguyen.

5.2.3 Documentation

This paper was written by Annabelle Austin. Editing was performed by all authors.

Statement of Authorship

Title of Paper	Patient selection for proton therapy: A radiobiological fuzzy Markov model incorporating robust plan analysis.
Publication Status	<input type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input checked="" type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Austin, A.M., Douglass, M.J.J., Nguyen, G.T. & Penfold, SN. Patient selection for proton therapy: A radiobiological fuzzy Markov model incorporating robust plan analysis. <i>Mathematical Medicine and Biology</i> . 2019 (Submitted May 2019)

Principal Author

Name of Principal Author (Candidate)	Annabelle Austin		
Contribution to the Paper	Developed methods and code, sourced input model parameters from literature performed simulations and analysis, wrote manuscript and acted as corresponding author.		
Overall percentage (%)	85%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature	_____	Date	27/05/19

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Scott Penfold		
Contribution to the Paper	Provided supervision, advice, and assisted in manuscript evaluation.		
Signature	_____	Date	22/05/19

Name of Co-Author	Michael Douglass		
Contribution to the Paper	Provided supervision, advice, and assisted in manuscript evaluation.		
Signature	_____	Date	22/05/19

Name of Co-Author	Giang Nguyen		
Contribution to the Paper	Provided supervision, advice, and assisted in manuscript evaluation.		
Signature		Date	24 / 05 / 2019

Patient selection for proton therapy: A radiobiological fuzzy Markov model incorporating robust plan analysis

Annabelle M. Austin^{a,d}, Michael J. J. Douglass^{a,b,e}, Giang T. Nguyen^{c,f},
Scott N. Penfold^{a,b,g}

^aDepartment of Physics, University of Adelaide,
Adelaide, SA 5005, Australia

^bDepartment of Medical Physics, Royal Adelaide Hospital
Adelaide, SA 5000, Australia

^cSchool of Mathematical Sciences, University of Adelaide,
Adelaide, SA 5005, Australia

^d Corresponding author email: annabelle.austin@adelaide.edu.au
Phone: +61 8 8313 5996
Fax: +61 8 8313 4380

^e Email: michael.douglass@adelaide.edu.au

^f Email: giang.nguyen@adelaide.edu.au

^g Email: Scott.Penfold@sa.gov.au

Abstract

Purpose: While proton therapy can offer increased sparing of healthy tissue compared with X-ray therapy, it can be difficult to predict whether a benefit can be expected for an individual patient. Predictive modelling may aid in this respect. However, the predictions of these models can be affected by uncertainties in radiobiological model parameters and in planned dose. The aim of this work is to present a Markov model that incorporates these uncertainties to compare clinical outcomes for individualised proton and X-ray therapy treatments.

Methods: A time-inhomogeneous fuzzy Markov model was developed which estimates the response of a patient to a given treatment plan in terms of quality adjusted life years. These are calculated using the dose-dependent probabilities of tumour control and toxicities as transition probabilities in the model. Dose-volume data representing multiple isotropic patient set-up uncertainties and range uncertainties (for proton therapy) are included to model dose delivery uncertainties.

Results: The model was retrospectively applied to an example patient as a demonstration. When uncertainty in the radiobiological model parameter was considered, the model predicted that proton therapy would result in an improved clinical outcome compared with X-ray therapy. However, when dose delivery uncertainty was included, there was no difference between the two treatments.

Conclusion: By incorporating uncertainties in the predictive modelling calculations, the fuzzy Markov concept was found to be well suited to providing a more holistic comparison of individualised treatment outcomes for proton and

X-ray therapy. This may prove to be useful in model-based patient selection strategies.

Keywords: Proton therapy, patient selection, Markov model, decision aid, radiobiological models

1 Introduction

As proton therapy becomes increasingly available, a greater number of clinics will need to decide which patients to treat with proton therapy instead of X-ray therapy. While a reduction in normal tissue complications can generally be expected with proton therapy, the expense and limited availability of the treatment suggests that the patients with the greatest need should be prioritised. Randomised Phase III clinical trial data comparing novel and standard treatments can become outdated in the rapidly evolving environment of radiation oncology. Hence this data may not always provide an adequate reference to determine which patients can expect the greatest benefit from receiving proton therapy. Alternatively, a modelling study in the form of an *in-silico* clinical trial [1] can be used to predict the patient outcome. There has been a growing interest in the clinical use of model-based methods to select patients for proton therapy [2,3].

In previous work, a Markov model was developed which used dosimetric data from a given treatment plan to determine probabilities of tumour control, second primary cancer induction, and multiple normal tissue complications [4]. The model combines these probabilities to give a single metric, the quality adjusted life expectancy (QALE), which allows a quantification of the effect of a treatment on a patient's quality of life. The result is a quantitative comparison of the clinical outcomes of two treatments on an individual patient basis. However, the limitation of the model was that the nominal planned dose was used in the comparison and there was no

consideration of the effect of dose and radiobiological model parameter uncertainties on the prediction.

The delivered dose can be different from the planned dose due to patient set-up errors or anatomical changes over the course of treatment. It has been suggested that these can have a significant impact on the accuracy of model-based selection of oropharyngeal cancer patients for proton therapy [5]. Range uncertainties are also an important consideration for proton therapy, where the distribution of dose with depth increases sharply at the end of the proton range. This feature is known as the Bragg peak and if this is misplaced, the result can be unacceptably high doses being delivered in normal tissue or an under-dosage of the clinical target volume (CTV). Furthermore, the parameters used in the radiobiological models are typically obtained with regression methods and are also subject to uncertainty. The combined impact of these uncertainties on the predictions of patient selection models warrants further investigation.

The aim of this work is to present an evolution of the patient selection model developed by Austin *et al.* [4], which includes a quantification of the effect of treatment and model uncertainties during model-based patient selection for proton therapy. The new model accounts for these sources of uncertainty using a fuzzy Markov model. Fuzzy Markov models are an extension of conventional Markov models where the transition probabilities are not known precisely [6]. The technical details of the model are presented in Section 2 with a description of the Markov states, transition probabilities and uncertainties. The model is demonstrated with a clinical example

(base of skull chordoma) in Section 3, and a discussion and conclusion are given in Section 4.

2 Methods

2.1 The Markov model

With a discrete-time inhomogeneous (time-dependent) Markov chain, the response of a patient to a particular treatment is modelled. Here, a revised version of the model developed by Austin *et al.* [4] is presented that takes the uncertainty of the input parameters into account. The reader is referred to Austin *et al.* [4] for the underlying details and motivation for the development of this model.

2.1.1 Markov states

A patient can occupy only a single Markov state at a given time. These states describe the health status of a patient and include Well, Deceased and the Unwell group of states. Figure 1 shows the Markov states for the simple case where only one injury with one grade is considered and multiple stages of cancer are not considered. In this figure, each state is represented by a node and allowed transitions are indicated by arrows between the relevant states.

In this work, the Unwell group contains states that represent varying numbers and forms of complications arising as a result of treatment. These states include the cases in which a patient:

- still has their initial primary cancer due to unsuccessful treatment, represented by the Cancer node;
- develops one or more normal tissue complications as a result of treatment, represented by the Inj node;
- develops one or more second primary cancers (SPCs) as a result of treatment, represented by the SPC node.

There are also other states representing every possible combination of the above:

- Unwell (primary cancer and injury): Cancer & Inj node;
- Unwell (primary and second cancer): Cancer & SPC node;
- Unwell (second cancer and injury): SPC & Inj node;
- Unwell (primary and second cancer and injury): Cancer & SPC & Inj node.

If there are multiple injuries being considered in the model, then a state exists for every possible combination of injuries. In addition, a patient can have any number of injuries as well as the primary cancer (if remaining after an unsuccessful treatment) and/or a second radiation-induced cancer.

The states with no injury present are duplicated and denoted by * in Figure 1. These are the states that a patient moves to after recovering from a particular injury. For each injury, there is a zero probability of moving from a recovered state back to the injured state, however a patient can develop a different injury from the recovered state. For example, if a patient develops only injury 1 and subsequently recovers,

they will be in the (Well, Injury 1*) state, where the star denotes a history of injury 1. From here it is possible to develop injury 2, but not injury 1. If injury 2 is developed and recovered from, then the patient will move to the (Well, Injury 1*, Injury 2*) state where the stars denote a history of both injuries. The recovered states were implemented to avoid a contradiction that would arise when a patient moved back to the Well state after recovering from an injury, only to develop it again some time later.

2.1.2 State transition probabilities

The time period modelled is the interval between the completion of the treatment and the death of the patient. The Markov chain consists of discrete time intervals, or cycles. Markov models in medical applications assume that a patient remains in a state for the duration of a cycle. The cycle length for this Markov chain was chosen to be one year, as this provided computational efficiency while maintaining sufficient temporal resolution. As a result, only toxicities occurring after one year are considered in the model. At the end of each cycle, it is possible for a patient to transition to another state. The time-dependent transition probabilities determine the likelihood of transitioning between particular state pairs at a given year of the Markov chain.

Many of the transition probabilities are derived from radiobiological models that take the planned dose to a structure as an input. These include:

- The tumour control probability (TCP), which is a single probability calculated using the dose-volume histogram (DVH) associated with the tumour volume [7].

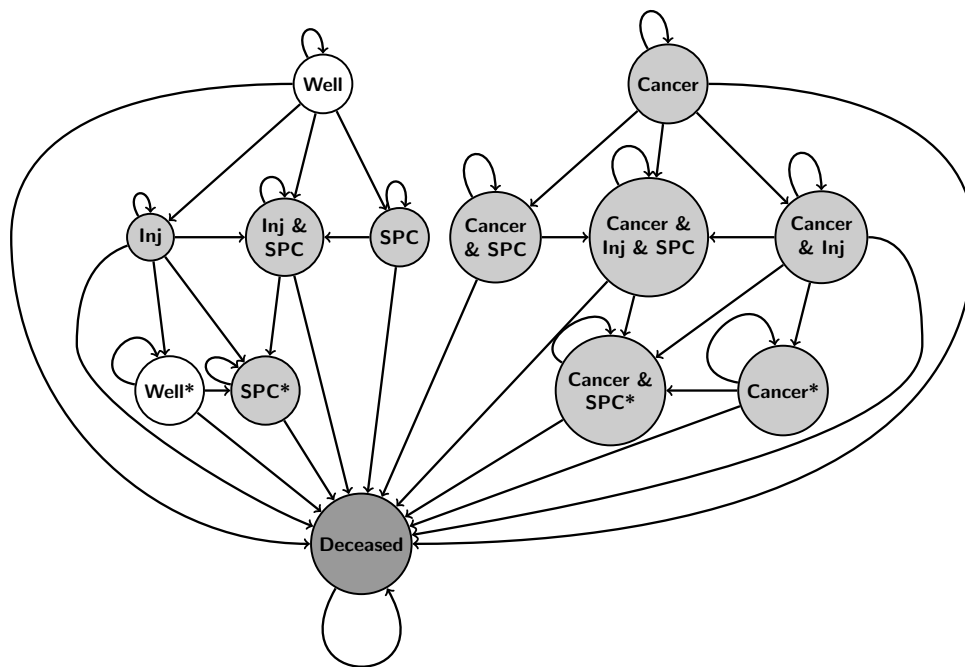


Figure 1: The Markov state transition diagram showing the allowed transitions between states. For simplicity, this describes the case where there is only one injury and one second primary cancer being considered in the model. ‘Well’ represents perfect health. ‘Cancer’ represents the situation where the patient still has the initial primary cancer. ‘SPC’ represents a state with a second primary cancer, and ‘Inj’ represents an injury state. The starred states denote a history of an injury. Unwell states are represented by pale grey nodes and the Deceased state is represented by a dark grey node.

Depending on the value of the TCP, the patient can begin the Markov chain in either the Well state or the Primary cancer state. Therefore, once the Markov chain has begun, it is not possible to transition between these two states, as depicted in Figure 1. The probability of beginning in the Well state is equal to the value of the TCP. The probability of beginning in the Primary cancer state is equal to the probability of the complement event, that is, 1 - TCP. It is not possible to begin the chain in any other state.

- The normal tissue complication probability (NTCP), which is the probability of developing a radiation-induced injury and is calculated using the DVH associated with a particular organ and the Lyman-Kutcher-Burman (LKB) model [8–10], given by

$$\text{NTCP} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t \exp\left(\frac{-x^2}{2}\right) dx, \quad (1)$$

$$\text{with } t = \frac{D_{eff} - TD_{50}}{mTD_{50}} \quad \text{and} \quad D_{eff} = \left(\sum_{i=1}^{\ell} v_i D_i^{\frac{1}{n}}\right)^n,$$

where TD_{50} is the uniform dose given to the entire organ that results in 50% complication risk, m is an organ specific parameter that is related to $\frac{d\text{NTCP}}{dD}$, n is a parameter that characterises the volume dependence of the organ's response to radiation, D_{eff} is the effective dose, and ℓ is the number of voxels. This is a time-dependent probability and the reader is referred to Austin *et al.* [4] for the details of the time-dependent probability calculation. Injury recovery was

not considered in this work. The injuries considered here include brainstem necrosis, tinnitus, blindness and endocrine dysfunction.

- The second primary cancer induction probability (SPCIP), which is the probability of developing a radiation-induced cancer as a result of the treatment. This is also a dose- and time-dependent quantity and was calculated using the model developed by Schneider *et al.* [11]. In this work, a single SPCIP is calculated using all contoured non-tumour volumes that represents the overall probability of having a second primary cancer in any structure.

The events of developing different injuries or a second cancer are assumed to be independent. Consequently, for the transitions that involve the development of several injuries and/or a second cancer simultaneously, the transition probability was assumed to be the multiplication of the relevant individual probabilities.

In addition to the dose-dependent transition probabilities, there are the probabilities of transitioning to the Deceased state. While these quantities are dose-independent and are not directly related to the treatment plan that is being compared, they allow for a more realistic estimate of the length and quality of life a patient can expect after receiving a given treatment. Uncertainties in these probabilities were not considered in this work. The probabilities of dying as a result of various causes include:

- Death from the primary cancer (with yearly probability 0.4, derived from 5-year survival rates after relapse of base of skull chordoma [12]) or a second primary cancer (with yearly probability 0.08 derived from cancer survival data [13]). These are time-independent quantities.

- Death from injury. This is also a time-independent quantity. In this work, it was assumed that it was not possible to die as a result of any injury, with the exception of brainstem necrosis which was assumed to be fatal.
- Unrelated death. This time-dependent probability is based on data from life tables obtained from the Australian Bureau of Statistics [14].

2.1.3 Quality of life utilities

Each state is assigned a quality of life (QoL) utility which represents the quality of life associated with the state, relative to perfect health. These utilities are used to calculate the QALE, which is the number of QALYs that are lived by a patient until they move to the absorbing Deceased state. A time horizon of 100 years was selected for the Markov model as the survival probability is negligible beyond this point. The QALE can be thought of as the number of years with a quality of life equivalent to perfect health lived by a patient after treatment. This is the primary metric used to evaluate and compare treatment plans in the proposed patient selection approach, as it incorporates both the probabilities of tumour control and of developing complications.

The QoL utilities applied in the Markov model are listed in Table 1. For Unwell states where there is more than one injury or cancer, the assigned utility is a multiplication of the utilities corresponding to the states where only one of each injury or cancer is present. The states representing the cases of second primary cancers were all assigned a value of 0.8 in accordance with the Eastern Cooperative Oncology Group (ECOG) performance status [15], as their definition of a grade 1 complication

QUALITY OF LIFE UTILITIES

State	Utility	Source	Comments
Skull base chordoma	0.72	[16]	
Brainstem necrosis	0.6	[15]	
Endocrine dysfunction	0.6	[15]	
Tinnitus	0.58	[17]	Evaluated after visiting a tinnitus clinic
Blindness	0.33	[18]	Complete blindness
Second primary cancer	0.8	[15]	

Table 1: Estimates for the quality of life (QoL) utilities for states in the Markov model.

(with utility 0.8) gives a reasonable description of these states. Where it was not possible to source an appropriate QoL utility for a particular injury, a default value of 0.6 was assigned as the ECOG definition of a grade 2 complication (which has a utility of 0.6) gives a reasonable description of these states.

2.1.4 Model evaluation

In the previous version of the model [4], the Markov chain was evaluated using a Monte Carlo approach. This involved simulating the Markov process many times with each simulation representing a different member of a hypothetical patient cohort. The disadvantage of this approach is that a large cohort is required for accurate results, corresponding to a significant computation time.

As an alternative, an analytical solution was implemented in this work. This enables the exact solution to be obtained with a significantly reduced computation time compared with the Monte Carlo method. Using the analytical approach, the distribution of a hypothetical cohort of patients – which are each an identical copy

of the real patient under consideration – amongst all Markov states, $\boldsymbol{\pi}$, is calculated. This is achieved with the use of a transition matrix, \mathbf{P} , which stores the probability of transitioning from each Markov state to all other Markov states. As this Markov model is time-dependent, the values of \mathbf{P} vary at each Markov cycle. Matrix multiplication is carried out to determine the distribution at a given time. The distribution of the cohort after x Markov cycles is given by

$$\boldsymbol{\pi}_x = \boldsymbol{\pi}_0 \prod_{k=0}^{x-1} \mathbf{P}_k, \quad (2)$$

where \mathbf{P}_k is the transition matrix corresponding to cycle k and $\boldsymbol{\pi}_0$ is the initial distribution of the cohort immediately after treatment (before any Markov cycles have taken place). Each element of the vector $\boldsymbol{\pi}$ gives the fraction of the cohort in a given state. Hence, all elements of $\boldsymbol{\pi}_0$ are zero except the elements representing the fraction of cohort who are well (equal to the TCP) and the fraction who did not have a successful treatment and still have the cancer (equal to $1 - \text{TCP}$).

The expectation value (or mean) of the life expectancy (LE), denoted $\widehat{\text{LE}}$, can be calculated with this data as

$$\widehat{\text{LE}} = \sum_{k=0}^m \sum_{j=1}^{n-1} (\boldsymbol{\pi}_k)_j \quad (3)$$

for n states and i years after treatment where the maximum value of i , $m = 100 - a$, corresponds to the year in which the patient (with age a at the time of treatment) is 100 years old. The Deceased state represented by the final element of $\boldsymbol{\pi}$ is not included in the summation.

To calculate the mean QALE, denoted $\widehat{\text{QALE}}$, each fraction is multiplied by the utility of the corresponding state:

$$\widehat{\text{QALE}} = \sum_{k=0}^m \sum_{j=1}^{n-1} (\boldsymbol{\pi}_k)_j \cdot u_j \quad (4)$$

where $\mathbf{U} = [u_1, \dots, u_{n-1}]$ is a vector of the utilities for each state.

2.2 Uncertainties

2.2.1 Radiobiological model parameters

The uncertainties in the radiobiological parameters manifest as uncertainties in the dose-dependent transition probabilities. The parameters in these models are usually determined with regression methods. In this work, uncertainties were only considered in the TD_{50} parameters for the pituitary dysfunction NTCP [19] and tinnitus [20], as the NTCPs associated with brainstem necrosis and blindness are typically small. The uncertainty was set to zero for other model parameters, but the framework exists to allow for non-zero uncertainties in these parameters.

2.2.2 Dose

The uncertainty related to the delivered dose was considered by performing analysis with the fuzzy numbers constructed from uncertainty in the radiobiological model parameters, for multiple scenarios of dose delivery variability. DVH curves corresponding to each scenario were calculated using the Varian Eclipse treatment planning system 13.7 (TPS) (Varian (Palo Alto, CA, USA)) and the relevant DVHs were

exported for use in the Markov chain. The result is a set of DVHs, one for each organ at risk (OAR) and the tumour volume for each scenario. Survival curves and QALEs were obtained for each scenario using the Markov chain.

2.3 The fuzzy Markov model

A ‘crisp’ Markov chain [6] can be thought of as a Markov chain where the values of the transition probabilities are known precisely. A fuzzy Markov chain does not have precisely known transition probabilities [6]: rather than a single probability X , each element of the transition matrix is an interval, $[\underline{X}, \overline{X}]$, where \underline{X} is the lower bound on the estimate of the value of X and \overline{X} is the upper bound. The upper and lower bounds are not necessarily symmetric about the crisp value in this model. In order to multiply fuzzy matrices, interval arithmetic must be applied. Matrix multiplication involves both multiplication and addition when considering elements; the rules for intervals are outlined as follows (following the notation of Moore *et al.* [21]):

- Addition

$$X + Y = [\underline{X} + \underline{Y}, \overline{X} + \overline{Y}]. \quad (5)$$

- Multiplication

$$X \cdot Y = [\min S, \max S], \text{ where } S = \{\underline{X}\underline{Y}, \underline{X}\overline{Y}, \overline{X}\underline{Y}, \overline{X}\overline{Y}\}. \quad (6)$$

The result is a confidence interval for each element of $\boldsymbol{\pi}$, with lower and upper bounds given by $\underline{\boldsymbol{\pi}}$ and $\overline{\boldsymbol{\pi}}$, which serves as a quantification of the uncertainty in the model output.

2.3.1 Optimization of the QALE uncertainty

Let the quality adjusted surviving fraction of the hypothetical cohort at a given time, $F(k)$, be defined as $F(k) = \sum_{j=0}^{n-n_A} (\boldsymbol{\pi}_k)_j \cdot u_j$, $k = [0, 1, 2, \dots, m]$, where there are n_A absorbing states. While $\sum_{j=0}^n \boldsymbol{\pi}_j = 1$ by definition, $\sum_{j=0}^n \underline{\boldsymbol{\pi}}_j$ and $\sum_{j=0}^n \overline{\boldsymbol{\pi}}_j$ may not necessarily be 1 for a given scenario of dose delivery accuracy. Therefore, a situation could arise where $\sum_{j=0}^n \overline{\boldsymbol{\pi}}_j > 1$, and therefore $\overline{F}(k)$ and $\underline{F}(k)$ cannot be defined as $\overline{F}(k) = \sum_{j=0}^{n-1} (\overline{\boldsymbol{\pi}}_k)_j$ and $\underline{F}(k) = \sum_{j=0}^{n-1} (\underline{\boldsymbol{\pi}}_k)_j$, respectively. As a result, defining $\overline{\text{QALE}}$ and $\underline{\text{QALE}}$ as $\overline{\text{QALE}} = \sum_{k=0}^m \overline{F}(k)$ and $\underline{\text{QALE}} = \sum_{k=0}^m \underline{F}(k)$, respectively, would be invalid.

As an alternative, the uncertainty in the QALE was determined through mathematical optimization. Using (2), it can be shown that the quality adjusted survival at a given year is a function of all the elements of the transition matrix \mathbf{P} , some of which are functions of the various NTCPs. Therefore, the QALE defined in (4) can be written as a function of the parameters in the radiobiological models as

$$\begin{aligned} f(\boldsymbol{\lambda}) = & f(p_{11}(\boldsymbol{\lambda}, m), \dots, p_{1n}(\boldsymbol{\lambda}, m), p_{21}(\boldsymbol{\lambda}, m), \dots, p_{nm}(\boldsymbol{\lambda}, x), \\ & p_{11}(\boldsymbol{\lambda}, m-1), \dots, p_{1n}(\boldsymbol{\lambda}, m-1), p_{21}(\boldsymbol{\lambda}, m-1), \dots, p_{nn}(\boldsymbol{\lambda}, m-1), \dots, \\ & p_{11}(\boldsymbol{\lambda}, 1), \dots, p_{1n}(\boldsymbol{\lambda}, 1), p_{21}(\boldsymbol{\lambda}, 1), \dots, p_{nn}(\boldsymbol{\lambda}, 1), \mathbf{U}), \end{aligned}$$

where m is the total number of Markov cycles, $\boldsymbol{\lambda} = [\lambda_1, \lambda_2] = [TD_{50t}, TD_{50e}]$, TD_{50t} is used to calculate the probability of tinnitus, and TD_{50e} is used to calculate the probability of endocrine dysfunction. Let $\underline{\lambda}_1$ and $\overline{\lambda}_1$ be the lower and upper bounds for TD_{50t} and $\underline{\lambda}_2$ and $\overline{\lambda}_2$ be the lower and upper bounds for TD_{50e} . The two optimization problems used to find the upper and lower bounds on the uncertainty have non-linear constraints and are defined as:

$$\begin{aligned}
& \min_{\boldsymbol{\lambda}} f(\boldsymbol{\lambda}) \\
\text{subject to } & \sum_{j=1}^n p_{ij}(k) = 1, \quad \text{for } i = 1, \dots, n \quad \text{and} \quad k = 1, \dots, m, \\
& \lambda_1 \in [\underline{\lambda}_1, \overline{\lambda}_1], \\
& \lambda_2 \in [\underline{\lambda}_2, \overline{\lambda}_2],
\end{aligned} \tag{7}$$

and

$$\begin{aligned}
& \max_{\boldsymbol{\lambda}} f(\boldsymbol{\lambda}) \\
\text{subject to } & \sum_{j=1}^n p_{ij}(k) = 1, \quad \text{for } i = 1, \dots, n \quad \text{and} \quad k = 1, \dots, m, \\
& \lambda_1 \in [\underline{\lambda}_1, \overline{\lambda}_1], \\
& \lambda_2 \in [\underline{\lambda}_2, \overline{\lambda}_2].
\end{aligned} \tag{8}$$

The constraints in the optimization problems ensure that each row of the transition matrix at each year must sum to 1. The objective function also depends on the quality of life utilities \mathbf{U} , which were not allowed to vary in the optimization problems.

The objective function and constraints were used with MATLAB's `fmincon()` routine for non-linear constrained optimization to determine the uncertainty bounds. To determine the uncertainty in the QALE as a result of radiobiological model parameter uncertainty only, the optimization routine was performed using the dose data for the nominal scenario. When considering dose delivery uncertainty in addition to the radiobiological model parameter uncertainty, the optimization was carried out for each dose delivery scenario d , and the upper and lower bounds are defined as:

$$\underline{\text{QALE}} = \min\{\underline{\text{QALE}}_1, \underline{\text{QALE}}_2, \dots, \underline{\text{QALE}}_D\},$$

and

$$\overline{\text{QALE}} = \max\{\overline{\text{QALE}}_1, \overline{\text{QALE}}_2, \dots, \overline{\text{QALE}}_D\}.$$

3 Demonstration patient

A four-year-old female presenting with base of skull chordoma was considered for the purposes of demonstrating the functionality of the model. Treatment plans for the delivery of 70 Gy in 35 fractions to the tumour with both protons (intensity modulated pencil beam scanning) and X-rays (volumetric modulated arc therapy (VMAT)) were retrospectively generated using the Varian Eclipse 13.7 TPS. In each treatment plan for this patient, each critical structure (healthy tissue) and the target volume corresponding to the tumour were contoured by a clinician. The VMAT plan was created on a planning target volume (PTV) with a 4 mm expansion of the CTV. The proton plan was generated through robust optimization of the CTV with a 3%

DOSE-DEPENDENT TRANSITION PROBABILITIES		
	Protons	VMAT
TCP	0.83 [0.64, 0.84]	0.74 [0.50, 0.83]
SPCIP	<0.01	0.020 [0.019, 0.076]
Brainstem necrosis NTCP	<0.01	<0.01
Tinnitus NTCP	0.05 [0.02, 0.12]	0.14 [0.08, 0.24]
Blindness NTCP	0.01 [<0.01, 0.05]	0.01 [<0.01, 0.05]
Endocrine dysfunction NTCP	0.99 [0.91, 1.0]	0.93 [0.56, 1.0]

Table 2: The values for the transition probabilities that were calculated and used as input for the Markov model, for each treatment modality. The SPCIPs listed represent the probability of a second cancer in the 25 years after treatment. Each probability calculated without considering uncertainties is listed along with its upper and lower bounds in square brackets when uncertainty is considered. TCP=tumour control probability.

range uncertainty and a 3 mm set-up uncertainty. A differential DVH was generated for each of the OARs and the CTV for both the nominal plan and the set of scenarios of dose delivery. These scenarios were produced in a robust plan analysis of both the VMAT and proton plan taking into account patient setup uncertainties and include +/- shifts of 3 mm in the (x, y, z) position of the patient with respect to isocentre (giving a total of 6 uncertainty scenarios). For proton therapy, there is an additional +/- beam range uncertainty of 3%, resulting in 12 scenarios.

The TCP was calculated based on dose to the CTV for both the proton and X-ray plans. The TCP, NTCPs and SPCIP that were calculated for each treatment plan for this patient are listed in Table 2.

The model returned the expectation value of the patient life expectancy along with the expectation value of the QALE. The results for this clinical example are listed in Table 3. The NTCP model parameter uncertainties do not contribute to the raw life expectancy, as it was assumed that the relevant injuries were not fatal.

ESTIMATED LIFE EXPECTANCIES

	Protons	VMAT
Raw LE	66.6 [53.4, 68.7]	57.9 [40.8, 65.7]
QALE (model uncertainties only)	39.5 [39.2, 40.0]	34.4 [33.0, 36.1]
QALE (model and dose uncertainties)	39.5 [31.4, 41.5]	34.4 [24.5, 45.8]

Table 3: The estimates of the mean raw life expectancies (LE) and the mean quality adjusted life expectancies (QALE) in years for each treatment modality. The results calculated without considering uncertainties are listed along with their respective upper and lower bounds in square brackets. The uncertainties in the QALE are listed with and without the inclusion of dose uncertainties.

Therefore, the uncertainties in the raw life expectancy arise from dose uncertainties only. The estimates of the life expectancy and QALE for the nominal proton plan are greater than the X-ray plan. When uncertainties in the radiobiological parameters only are considered, there would be a clear benefit for the patient if treated with proton therapy. However, when the uncertainties associated with treatment delivery are accounted for, the proton plan does not have a significantly increased benefit compared with the X-ray plan. The survival curves are given in Figures 2 and 3. The uncertainty in the quality adjusted survival curve was calculated with (2) and the maximum and minimum values of TD_{50} that satisfied the constraints of the optimization problem. It is apparent that uncertainty in the NTCP parameters has a smaller effect than the set-up and range uncertainties, and that the uncertainty associated with proton therapy is smaller than that of VMAT.

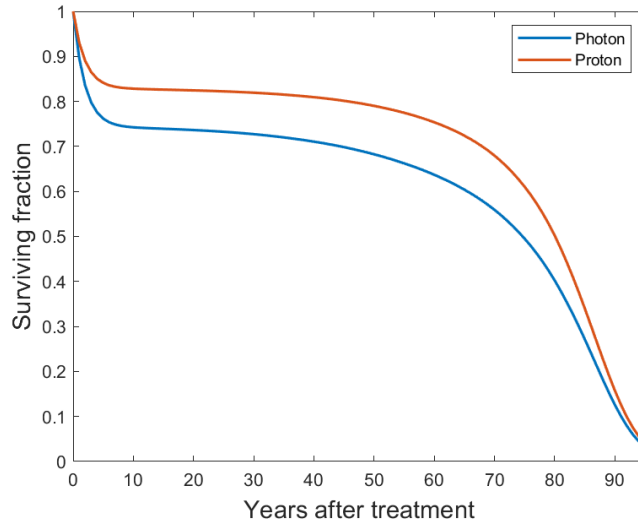


Figure 2: Raw survival curves for both the proton plan and VMAT plan.

4 Discussion

For the presented example, the uncertainty associated with proton therapy was predicted to be smaller than the uncertainty associated with VMAT. This is the case for both the dose and radiobiological model parameter uncertainties. The later is due to the differences in the effective dose, D_{eff} between the two treatments, resulting in different t parameters. The implication of a t parameter corresponding to a steeper part of the dose response curve will be a greater magnitude in NTCP uncertainty if t is varied within a certain range. The difference in the magnitude of the dose uncertainty is likely a result of the proton plan being robustly optimized to the CTV and the photon plan being robustly optimized to the PTV.

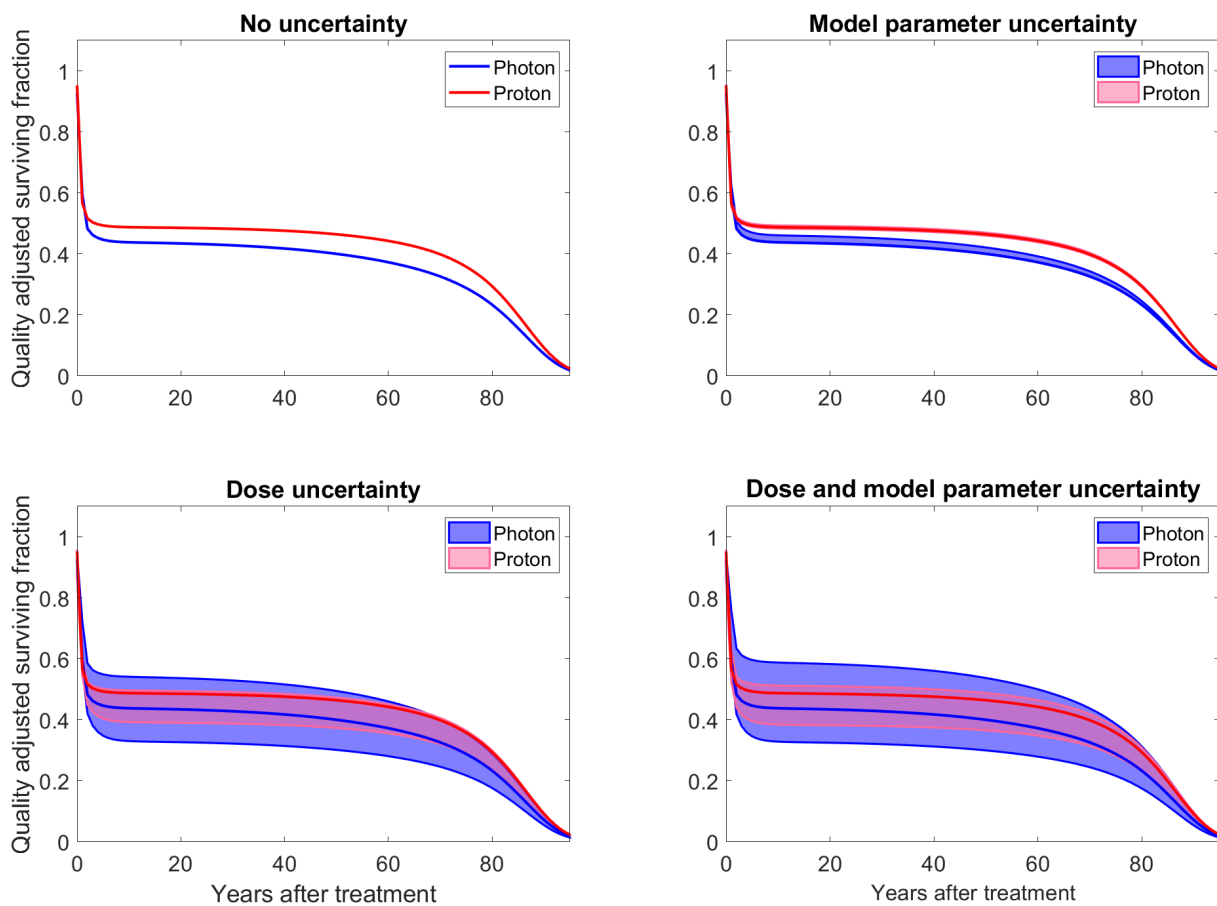


Figure 3: Quality adjusted survival curves for both the proton plan and the VMAT plan. The top right figure corresponds to a situation in which only uncertainty in NTCP model parameters is considered. The bottom left figure corresponds to a situation in which only uncertainty in dose is considered. The bottom right figure corresponds to the situation in which both dose and NTCP model parameter uncertainties are included.

The presented model improves on the existing model [4] by considering the effects of dose delivery and radiobiological model parameter uncertainties associated with a radiotherapy treatment on whether an individual patient would be selected for proton therapy. However, the model has several limitations. It should be stressed that not all transition probabilities and quality of life utilities in the presented example were clinically founded, as it was not always possible to obtain representative values in the literature. As in any predictive model, the usefulness of the results is directly dependent on the accuracy of the input parameters. At this stage, the functionality of the model has been demonstrated with an example patient and appropriate validation of all radiobiological model parameters must be conducted before clinical implementation of this model-based patient selection approach. *In-silico* clinical trials could potentially be used to gradually refine the accuracy of the input data over time through a feedback system [1].

However, in some cases, it is not possible to know the true transition probability for a given state pair. This is particularly true for patient death in the cases where there are multiple possible causes of death. For example, for the fraction of the cohort in a state where there are multiple cancers or fatal injuries, there are different ways of transitioning to the absorbing state. However, the death probabilities used in the model correspond to the probability of transitioning from a state representing a given complication to the Deceased state, and do not reflect the probability of dying when multiple complications are present. In these cases, it was assumed that the probability of death was the maximum death probability out of all possible causes. The result is that patients with multiple complications have the same probability

of dying compared to patients with a single complication. This is a reasonable approximation, as there is usually a single dominant death probability and it is unlikely that a patient could have a fatal cancer or injury while simultaneously having an equally high background death probability.

Several of the parameters in the radiobiological models used in this study were determined in studies based on adult populations. However, the output of the Markov model is demonstrated using a paediatric patient in this work. It is possible that this assumption contributed to uncertainty in the results, as it is likely that younger cell populations have differing radiobiological properties compared with older populations.

The estimated SPCIPs for the example patient were lower compared with observations of second malignancy incidences. These have been found to be 7.5% for photons and 5.3% for protons (median follow up of 6.7 years) [22]. The model used to calculate the SPCIP in this work is yet to be validated, and this is the likely reason for the discrepancy between this work and the observations.

The dose uncertainty included in the model is derived from multiple scenarios of treatment delivery accounting for variations in delivered dose as a result of patient positioning and range uncertainties. These include worst case scenarios modelled with a systematic set-up uncertainty applied throughout the treatment course, which is unlikely to occur in reality. The result is an overestimation of the magnitude of the uncertainty in the final results. In future applications of the model, the robust analysis will be refined to allow the generation of more realistic scenarios through the incorporation of fractionation effects of random set-up uncertainties [23].

As the dose distribution will influence tumour control rates and complication rates, it is entirely possible that uncertainties in the radiobiological parameters naturally incorporate the effect of uncertainties in the dose delivery. As we consider dose uncertainties in addition to uncertainties in these parameters, there is the possibility that the quoted uncertainties in the QALE are overestimated due to the doubling of the dose uncertainty. However, the advantage of considering the dose uncertainties separately is that it allows the contribution of these uncertainties to be directly assessed and isolated.

4.1 Conclusion

The presented model could serve as a valuable tool for patient selection for proton therapy. The effects of radiobiological model parameter and dose uncertainties have been included to aid decision making in the referral process. In this work, the functionality of the model has been demonstrated. The inclusion of the uncertainties demonstrated the need for validated and precise radiobiological model parameters in model-based patient selection strategies.

Acknowledgements. The authors wish to thank Peter Rhodes for the early development of the project, Alexandre Santos for the development of the SPCIP code and Raymond Dalfsen for producing the proton and photon patient plans used in this study. The first author acknowledges the support of an Australian Government Research Training Program (RTP) Scholarship. The third author acknowledges the support of ACEMS (ARC Centre of Excellence for Mathematical and Statistical Frontiers).

Compliance with ethical standards

Conflict of Interest: Scott Penfold has worked part-time for a developer of a proton therapy centre. The other authors declare that they have no conflict of interest.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee (Royal Adelaide Hospital Research Ethics Committee No. 150322) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

References

- [1] Langendijk JA, Lambin P, De Ruyscher D, et al. Selection of patients for radiotherapy with protons aiming at reduction of side effects: the model-based approach. *Radiotherapy and Oncology*. 2013;107(3):267–273.
- [2] Blanchard P, Wong AJ, Gunn GB, et al. Toward a model-based patient selection strategy for proton therapy: external validation of photon-derived normal tissue complication probability models in a head and neck proton therapy cohort. *Radiotherapy and Oncology*. 2016;121(3):381–386.
- [3] Jakobi A, Bandurska-Luque A, Stützer K, et al. Identification of patient benefit from proton therapy for advanced head and neck cancer patients based on individual and subgroup normal tissue complication probability analysis. *International Journal of Radiation Oncology* Biology* Physics*. 2015;92(5):1165–1174.
- [4] Austin AM, Douglass MJ, Nguyen GT, et al. A radiobiological markov simulation tool for aiding decision making in proton therapy referral. *Physica Medica*. 2017;44:72–82.
- [5] Bijman RG, Breedveld S, Arts T, et al. Impact of model and dose uncertainty on model-based selection of oropharyngeal cancer patients for proton therapy. *Acta Oncologica*. 2017;56(11):1444–1450.
- [6] Buckley JJ. *Fuzzy probabilities: New approach and applications*. Vol. 115. Springer Science & Business Media; 2005.

- [7] Allen Li X, Alber M, Deasy JO, et al. The use and QA of biologically related models for treatment planning: Short report of the TG-166 of the therapy physics committee of the AAPM. *Medical Physics*. 2012;39(3):1386–1409.
- [8] Lyman JT. Complication probability as assessed from dose-volume histograms. *Radiation Research*. 1985;104(2s):S13–S19.
- [9] Kutcher GJ, Burman C. Calculation of complication probability factors for non-uniform normal tissue irradiation: The effective volume method. *International Journal of Radiation Oncology* Biology* Physics*. 1989;16(6):1623–1630.
- [10] Burman C, Kutcher G, Emami B, et al. Fitting of normal tissue tolerance data to an analytic function. *International Journal of Radiation Oncology Biology Physics*. 1991;21(1):123–135.
- [11] Schneider U, Sumila M, Robotka J. Site-specific dose-response relationships for cancer induction from the combined Japanese A-bomb and Hodgkin cohorts for doses relevant to radiotherapy. *Theoretical Biology and Medical Modelling*. 2011;8(1):1–21.
- [12] Fagundes MA, Hug EB, Liebsch NJ, et al. Radiation therapy for chordomas of the base of skull and cervical spine: patterns of failure and outcome after relapse. *International Journal of Radiation Oncology* Biology* Physics*. 1995; 33(3):579–584.

- [13] Australian Institute of Health and Welfare Australian Cancer Incidence and Mortality books: all cancers combined Canberra: AIHW ; 2017. Available from: <http://www.aihw.gov.au/acim-books/>, Accessed: 3-11-2018.
- [14] Life Tables, States, Territories and Australia, 2013-2015 ; ???? Available from: <http://http://www.abs.gov.au/ausstats/abs@.nsf/mf/3302.0.55.001>, Accessed: 17-01-2019.
- [15] Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *American Journal of Clinical Oncology*. 1982;5(6):649–656.
- [16] Ramaekers BL, Grutters JP, Pijls-Johannesma M, et al. Protons in head-and-neck cancer: bridging the gap of evidence. *International Journal of Radiation Oncology* Biology* Physics*. 2013;85(5):1282–1288.
- [17] Refaie AE, Davis A, Kayan A, et al. A questionnaire study of the quality of life and quality of family life of individuals complaining of tinnitus pre-and postattendance at a tinnitus clinic. *International Journal of Audiology*. 2004; 43(7):410–416.
- [18] Tengs TO, Wallace A. One thousand health-related quality-of-life estimates. *Medical Care*. 2000;38(6):583–637.
- [19] De Marzi L, Feuvret L, Boulé T, et al. Use of gEUD for predicting ear and pituitary gland damage following proton and photon radiation therapy. *The British Journal of Radiology*. 2015;88(1048):20140413.

- [20] Lee TF, Yeh SA, Chao PJ, et al. Normal tissue complication probability modeling for cochlea constraints to avoid causing tinnitus after head-and-neck intensity-modulated radiation therapy. *Radiation Oncology*. 2015;10(1):194.
- [21] Moore RE, Kearfott RB, Cloud MJ. Introduction to interval analysis. Vol. 110. Siam; 2009.
- [22] Chung CS, Yock TI, Nelson K, et al. Incidence of second malignancies among patients treated with proton versus photon radiation. *International Journal of Radiation Oncology* Biology* Physics*. 2013;87(1):46–52.
- [23] Lowe M, Albertini F, Aitkenhead A, et al. Incorporating the effect of fractionation in the evaluation of proton plan robustness to setup errors. *Physics in Medicine & Biology*. 2015;61(1):413.

5.3 Fuzzy Markov model details

The development of the Fuzzy model involved assigning uncertainties to the transition probabilities in the Markov model. The only probabilities that did not have uncertainty were the background death probabilities. Error bounds were assumed to be asymmetric.

The uncertainties in the radiobiological model parameters and the input dose data have been discussed in the publication P2. There are additional uncertainties included in the model that were not considered in the publication. The primary reason for this was due to a lack of available data on uncertainties. The quality of life utilities have a significant influence on the QALE predicted for a given patient and treatment plan. Uncertainties in these weights were incorporated into the Fuzzy model. In addition, while the publication P2 focussed on uncertainties in the dose-dependent transition probabilities, there was no consideration of dose-independent transition probabilities. These include death probabilities, injury recovery probabilities and the probabilities of transitioning to higher stages of cancer. The Fuzzy model includes uncertainties in these probabilities, with the uncertainties being easily modifiable by the user if reliable parameters become available.

5.3.1 NTCP calculation

The NTCP was calculated for each injury using (2.8) (Section 2.3.2). The parameters used in the NTCP calculations are listed in Table 5.1.

TABLE 5.1: Parameters used for the calculation of the all-time NTCP for each injury using the LKB model. Burman *et al.* [62] was used as the source of the parameters unless otherwise stated. Estimates of the mean (\bar{x}) and standard deviation (σ_x) of the time taken for each injury to develop after treatment are also listed in months. Where it was not possible to find an estimate of the mean or standard deviation, it was assumed to be 12 months or 6 months, respectively. Confidence intervals are indicated in square brackets for the TD_{50} parameters where applicable.

Tissue	α/β (Gy)	Endpoint	n	m	TD_{50} (Gy)	\bar{x}	σ_x
Brainstem	2.5 ¹	Necrosis	0.16	0.14	65	17 ²	6 ³
Ear	3.0 ⁴	Tinnitus	0.01	0.35 ⁵	46.5 [41.9, 53.4] ⁵	12 ³	6 ³
Optic chiasm	3.0 ⁴	Blindness	0.25	0.14	65	18 ⁶	6 ³
Pituitary	3.0 ⁴	Endocrine dysfunction	0.156 ⁷	0.08 ⁷	60.6 [59.1, 62.0] ⁷	12 ³	6 ³

¹ Source: [86]

² Source: [87]

³ Not clinically founded

⁴ Assumed as the default 3.0 for late responding tissue

⁵ Source: [64]

⁶ Source: [88]

⁷ Source: [63]

5.3.2 Interval arithmetic

As discussed in the publication P2 Section 2.2, interval arithmetic is applied to evaluate the Fuzzy model.

Following the notation used in Moore [89], an interval X can be represented as $X = [\underline{X}, \overline{X}]$. Simple arithmetic operations can be carried out through applying the following rules:

- Addition

$$X + Y = [\underline{X} + \underline{Y}, \overline{X} + \overline{Y}]. \quad (5.1)$$

- Subtraction

$$X - Y = [\underline{X} - \bar{Y}, \bar{X} - \underline{Y}]. \quad (5.2)$$

- Multiplication

$$XY = [\min S, \max S], \quad \text{where } S = \{\underline{XY}, \underline{X}\bar{Y}, \bar{X}\underline{Y}, \bar{X}\bar{Y}\}. \quad (5.3)$$

- Division

$$X/Y = X(1/Y) \quad \text{where } 1/Y = [1/\bar{Y}, 1/\underline{Y}]. \quad (5.4)$$

These operations allow fuzzy transition matrices to be multiplied.

5.3.3 Quality of life utility weights

The quality of life utility weights are a major contributor to the QALE estimated from a given plan, which is the primary metric used to evaluate the plan quality (Section 2.2.3). It can be difficult to quantify the impact of an illness on a patient's quality of life, relative to perfect health. As a result, estimates of quality of life utilities often include an uncertainty estimate [90].

The model presented in this chapter treats the quality of life utilities as fuzzy numbers. The interval arithmetic rules defined in Section 5.3.2 are applied when multiplying the vector of utilities \mathbf{U} with $\boldsymbol{\pi}$ (in (3) of the publication P2) to calculate the bounds on the QALE. While it was not possible to source uncertainties for utilities for the injuries and cancers considered in P2, the framework exists in the model to include uncertainties if required. The utilities and associated uncertainties may be modified by the user in this instance.

The lower bounds on the utilities for states where there are more than one injury or cancer present is the multiplication of the lower bounds on the utilities of the individual injuries or cancers. This method is also applied to calculate the upper bounds.

5.3.4 Transition matrix

Once all of the individual transition probability confidence intervals (CIs) are calculated, they are used to calculate the upper and lower bounds on each element of the transition matrix \mathbf{P} . The CIs of all matrix elements are calculated using the lower bounds on the CIs of the each of the relevant NTCPs, SPCIPs, death and recovery probabilities. The same is true for the upper bounds. An important implication is that the lower bounds on the CIs of all elements of a given row will not necessarily sum to one. This also applies to the upper bounds.

Organs with negligible NTCPs are not considered in the model (see Section 4.2.4). However, an organ may have a negligible NTCP in one scenario of dose delivery accuracy, while having a non-negligible NTCP in another. It was therefore necessary to recalculate the number of injuries and the number of states for each scenario. This process allows the model to run faster for some scenarios, and hence the model as a whole can be more efficient.

For the transitions between absorbing states, the diagonal elements are $[1, 1]$ and the off-diagonal elements are $[0, 0]$, as there are no uncertainties associated with these.

5.4 Discussion and conclusion

The Fuzzy model has been developed to include the effect of several sources of uncertainty on the model predictions. Optimizing the QALE uncertainty demonstrated that the patient considered could expect a clinical benefit as a result of proton therapy (in the presence of NTCP model parameter uncertainty only). Including more realistic scenarios of dose delivery accuracy in treatment planning systems could reduce the magnitude of uncertainties in dose delivery. A challenge in developing this model was sourcing appropriate input uncertainties, particularly for the quality of life utility weights. The model presented in this chapter contains the framework to analyse these uncertainties, once they become available. Ultimately, comparing the model predictions with patient outcomes through clinical validation will allow the estimated uncertainties to be assessed.

Chapter 6

Cost-effectiveness Model

The publication P3 forms the basis of this chapter:

Austin, A.M., Douglass, M.J.J., Nguyen, G.T., Dalfsen, R., Le, H., Gorayski, P., Tee, H., Penniment, M. & Penfold, S.N.. Cost-effectiveness of proton therapy in treating base of skull chordoma. *Australasian Physical and Engineering Sciences in Medicine*. 2019 (Submitted May 2019).

6.1 Introduction and motivation

In Chapter 5, the incorporation of model parameter uncertainties into the AE model is described. However, this model did not consider the effects of treatment cost-effectiveness. This is particularly important when considering proton therapy, where treatment costs are significantly greater when compared with X-ray therapy. Treatment costs are also an important consideration when developing health policy more generally.

In this chapter, the method of incorporating cost-effectiveness into the AE model (at the stage before uncertainties were incorporated) is described. The motivation of this work was to develop the functionality of the patient selection model to make decisions based on treatment cost-effectiveness. In some cases where there is an elevated risk of treatment failure or radiation-induced injury or cancer associated with an X-ray treatment, the

larger initial cost of proton therapy may be justified if it is less than the combined cost of X-ray therapy and of treating complications.

The aim of the publication that is the basis of this chapter was to present the model with cost-effectiveness incorporated and utilise a cohort of base of skull chordoma patients to determine whether this indication may be treated with proton therapy cost-effectively. The advantage of presenting a cohort rather than an individual patient is that it enables a more powerful conclusion to be drawn regarding the cost-effectiveness of the treatment.

6.2 Statement of contribution

6.2.1 Conception

The idea of incorporating cost-effectiveness into the model was first conceptualised by Scott Penfold. Annabelle Austin developed the approach of incorporating cost-effectiveness.

6.2.2 Realisation

The writing of the code, analysis, and sourcing of parameters was performed by Annabelle Austin. The organs on the CT scans were contoured by Hien Le, Peter Gorayski, Hui Tee and Michael Penniment. The radiotherapy treatment plans used in the analysis were created by Raymond Dalfsen.

6.2.3 Documentation

This paper was written by Annabelle Austin. Editing was performed by all authors.

Statement of Authorship

Title of Paper	Cost-effectiveness of proton therapy in treating base of skull chordoma.
Publication Status	<input type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input checked="" type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Austin, A.M., Douglass, M.J.J., Nguyen, G.T., Dalfsen, R., Le, H., Gorayski, P., Tee, H., Penniment, M. & Penfold, S.N.. Cost-effectiveness of proton therapy in treating base of skull chordoma. <i>Australasian Physical and Engineering Sciences in Medicine</i> . 2019 (Submitted May 2019).

Principal Author

Name of Principal Author (Candidate)	Annabelle Austin		
Contribution to the Paper	Developed methods and code, sourced input model parameters from literature performed simulations and analysis, wrote manuscript and acted as corresponding author.		
Overall percentage (%)	85%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	28/05/19

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Scott Penfold		
Contribution to the Paper	Provided supervision, advice, and assisted in manuscript evaluation.		
Signature		Date	22/05/19

Name of Co-Author	Michael Douglass		
Contribution to the Paper	Provided supervision, advice, and assisted in manuscript evaluation.		
Signature		Date	22/05/19

Name of Co-Author	Giang Nguyen		
Contribution to the Paper	Provided supervision, advice, and assisted in manuscript evaluation.		
Signature		Date	24/05/2019

Name of Co-Author	Raymond Daltsen		
Contribution to the Paper	Created treatment plans used in analysis and assisted in manuscript evaluation.		
Signature		Date	24/05/2019

Name of Co-Author	Hien Le		
Contribution to the Paper	Performed organ contouring to facilitate creation of treatment plans used in analysis and assisted in manuscript evaluation.		
Signature		Date	24/5/19

Name of Co-Author	Peter Gorayski		
Contribution to the Paper	Performed organ contouring to facilitate creation of treatment plans used in analysis and assisted in manuscript evaluation.		
Signature		Date	28/05/19

Name of Co-Author	Hui Tee		
Contribution to the Paper	Performed organ contouring to facilitate creation of treatment plans used in analysis and assisted in manuscript evaluation.		
Signature		Date	28/05/19

Name of Co-Author	Michael Penniment		
Contribution to the Paper	Performed organ contouring to facilitate creation of treatment plans used in analysis and assisted in manuscript evaluation.		
Signature		Date	25/5/19

Cost-effectiveness of proton therapy in treating base of skull chordoma

Running title: Proton therapy cost-effectiveness

Authors:

Annabelle M. Austin¹

Michael J. J. Douglass^{1,2}

Giang T. Nguyen³

Raymond Dalftsens⁴

Hien Le⁴

Peter Gorayski⁴

Hui Tee⁴

Michael Penniment⁴

Scott N. Penfold^{1,2}

¹Department of Physics, University of Adelaide, Adelaide, SA, Australia

²Department of Medical Physics, Royal Adelaide Hospital, Adelaide, SA, Australia

³School of Mathematical Sciences, University of Adelaide, Adelaide, SA, Australia

⁴Department of Radiation Oncology, Royal Adelaide Hospital, Adelaide, SA, Australia

Corresponding author:

Annabelle M. Austin

Department of Physics

School of Physical Sciences

University of Adelaide

North Terrace

Adelaide, 5005

SA, Australia

Email: annabelle.austin@adelaide.edu.au

Abstract

Introduction: While proton beam therapy (PBT) can offer increased sparing of healthy tissue, it is associated with large capital costs and as such, has limited availability. Furthermore, it has not been well established whether PBT has significant clinical advantages over conventional volumetric modulated arc therapy (VMAT) for all tumour types. PBT can potentially offer improved clinical outcomes for base of skull chordoma (BOSCh) patients compared with photon (X-ray) therapy, however the cost-effectiveness of these treatments is unclear. In this study, the cost-effectiveness of PBT in the treatment of BOSCh patients is assessed, based on an analysis of comparative radiotherapy treatment plans using a radiobiological Markov model.

Methods: Seven BOSCh patients had treatment plans for the delivery of intensity modulated proton therapy (IMPT) and VMAT retrospectively analysed. The patient outcome (in terms of tumour local control and normal tissue complications) after receiving each treatment was estimated with a radiobiological Markov model. In addition, the model estimated the cost of both the primary treatment and treating any resultant adverse events. The incremental cost-effectiveness ratio (ICER) was obtained for each patient.

Results: PBT was found to be cost-effective for 6 patients and cost-saving for 1. The mean ICER was AUD\$3,220 per quality adjusted life year (QALY) gained. Variation of model parameters resulted in the proton treatments remaining cost-effective for these patients.

Conclusion: Based on this cohort, PBT is a cost-effective treatment for patients with BOSCh. This supports the inclusion of PBT for BOSCh in the Medicare Services Advisory Committee 1455 application.

Keywords

Proton therapy, cost-effectiveness, base of skull chordoma, Markov model, radiobiological models

1 Introduction

Medicare support for proton beam therapy (PBT) in Australia is currently being considered as part of the Medicare Services Advisory Committee (MSAC) 1455 application. MSAC 1455 considers PBT for a specific list of cancer types and has included a review of clinical evidence for PBT for these cancers. The assessment identified a lack of Level 1 evidence for PBT across multiple tumour types. Due to issues regarding equipoise, funding and availability, there have not been any Phase III randomized clinical trials comparing PBT to conventional photon (X-ray) therapy for the cancer types listed in MSAC 1455. In this case, a lack of evidence does not equate to non-superiority. Therefore, it is important that other approaches are considered when assessing whether new technologies should be supported for funding through the public health system.

Markov models were adopted by the Assessment of New Radiation Oncology Technology and Treatments (ANROTAT) project¹, undertaken by the Trans-Tasman Radiation Oncology Group (TROG) and funded by the Australian Federal Government Department of Health and Aging. The group recommends that Markov models be adopted for economic assessments of new health technologies. In the current work, we propose the use of Markov models for assessing the cost-effectiveness of PBT relative to conventional X-ray therapy.

One of the most common indications making use of the Medical Treatment Overseas Program (funded by the Australian Government Department of Health) for PBT is base of skull chordoma (BOSCh). Chordoma is a very rare form of bone tumour, accounting for 1–4% of all primary malignant bone tumours² with base of skull cases representing approximately one third of presentations³. Achieving complete surgical removal can be limited by the critical anatomical location. Similarly, postoperative radiotherapy with X-rays can also be limited by the presence of nearby critical organs.

Mailhot Vega *et al.*⁴ have proposed a method of selecting paediatric brain cancer patients to receive PBT based on treatment cost-effectiveness. PBT was found to be cost-effective or even cost-saving, depending on the degree to which the hypothalamus could be spared with protons compared with photons. Peeters *et al.*⁵ have carried out a cost analysis of treating various indications with particle therapies compared with photon therapy, based on construction and operational costs. Treatment costs for various tumour types were sourced through a review of cost-effectiveness studies. Cost differences between particle and photon therapy were found to be larger for BOSCh treatments compared with lung and prostate treatments. The cost-effectiveness of carbon ion therapy in the treatment of BOSCh has been analysed by Jäkel *et al.*⁶, based on studies of local control improvement compared with photon therapy. Primary treatment costs and costs for recurrent tumours were estimated. It was found that if local control exceeds 70% with carbon ion therapy, then the overall treatment costs of carbon ion therapy are lower than that of conventional radiotherapy (assuming a local control rate of 50%). The limitation of their approach is that costs associated with toxicities and productivity losses were not considered. Therefore, it may be possible that carbon ion therapy is cost-effective at a smaller difference in local control. Lundkvist *et al.*⁷ have included the effects of adverse events to evaluate the cost-effectiveness of PBT in the treatment of childhood medulloblastoma. A Markov simulation model was used to combine risks of a wide range of toxicities including hearing loss, intelligence quotient (IQ) loss, hypothyroidism, growth hormone deficiency (GHD), osteoporosis, cardiac disease, and second malignancies. PBT was found to be cost-effective and cost-saving compared with conventional radiation therapy for patients with a high risk of IQ loss or developing GHD. However, in this approach variations in the dosimetry

between individual patient treatment plans was not considered directly (population-based risks were applied).

In previous work by our group, a Markov model was developed with the ability to identify patients who would receive the most improved clinical outcome if treated with PBT compared with X-ray therapy⁸. The model predicts the radiobiological effect of a given treatment plan on an individual patient basis. This effect includes contributions from locoregional control, treatment toxicities and radiation-induced malignancies. The inclusion of second radiation-induced cancer risk is particularly important when considering younger patients (who comprise the majority of BOSCh patients) as they have a longer remaining life-time over which to develop second cancers. The output of the Markov model was the quality-adjusted life expectancy (QALE), or the number of quality adjusted life years (QALYs) associated with a radiotherapy treatment plan. This output allows quantitative comparisons of treatment modalities.

In the current work, the previously developed Markov model is extended to include a cost-effectiveness analysis, with the output being the cost of a treatment per QALY gained, also known as the incremental cost-effectiveness ratio (ICER). This work builds on that of Mailhot Vega *et al.*⁴ with the inclusion of second cancer risk, locoregional control and a wider range of potential radiation-induced injuries. The aim of this work was to determine whether BOSCh patients can be treated with PBT cost-effectively, based on individual patient dosimetric analyses.

2 Methods

2.1 Patient cohort and treatment planning

The cohort consisted of 7 female BOSCh patients with a wide range of ages at the time of treatment. The size and characteristics of the cohort was limited by availability as BOSCh is particularly rare. The ages and prescription doses are summarised in Table 1. Each patient had volumetric modulated arc therapy (VMAT) and intensity modulated proton therapy (IMPT) treatment plans generated in the Varian Eclipse treatment planning system. VMAT plans consisted of 2 co-planar arcs using a 6 MV Varian TrueBeam HD MLC beam model clinically commissioned at the Royal Adelaide Hospital (RAH). Plans were optimized to a planning target volume (PTV) which was generated from a 3 mm expansion of the CTV. IMPT plans consisted of 2-4 beams with pencil beam weights obtained through robust multi-field optimization (MFO) to the clinical target volume (CTV). Beam range uncertainty of 3% and set-up uncertainties of +/-3mm were included in the robust optimization. The proton beam model was based on a Varian ProBeam accelerator.

Table 1: The patient ages at the time of treatment and prescription doses.

Patient ID	Age (years)	Treatment schedule (Gy/fraction #)	Comments
1	6	78/39	
2	12	78/39	
3	8	78/39	
4	46	70/35	
5	27	74/37	
6	51	74/37	Pituitary not discernible
7	4	70/35	CT scan did not extent to parotids

2.2 Markov model

A discrete-time Markov chain model developed previously⁸ was extended in this work to model the clinical outcome of each patient. The model consists of several Markov states, with each representing a unique status of health. These include the Well state (or

complication-free control), the Deceased state, and the states representing various treatment complications. These are detailed in Sections 2.4 and 2.5. In addition, there are states representing an unsuccessful treatment.

It is assumed that a patient occupies a single state at a given time. In each cycle (defined to be one year in this work), it is possible for the patient to transition to another state based on certain probabilities. For example, the probability of transitioning from the Well state to an injured state is given by the calculated normal tissue complication probability (NTCP) corresponding to the injury for the patient being considered.

The transition probabilities in the model can be either dose-dependent or dose-independent. The probabilities of locoregional control, second cancer induction and normal tissue complication are dose-dependent and are calculated using the dose-volume histogram (DVH) data from a given treatment plan. Death and recovery probabilities were assumed to be dose-independent in this work.

2.3 Locoregional control

The DVH data for the tumour volume for each patient was used to determine a tumour control probability (TCP) that was unique for each patient (details described by Austin *et al.*⁸). In the event of treatment failure, it is assumed the patients cannot return to the well state (i.e. no retreatments). While this is a simplification of the disease progression, the clinical outcome of the two alternate primary treatments are the subjects of comparison in the model.

The yearly death probability (due to treatment failure), denoted $\Pr(Die)$, applied in this analysis was 0.4. This was derived from 5-year survival rates after relapse (local or distant) of base of skull chordoma⁹ (7%), by evaluating $S = (1 - \Pr(Die))^n$, where $n=5$ and $S=0.07$. Solving for the death probability gives 0.4. There is an additional risk of death each year due to unrelated causes. The annual probability of this was derived using Life Tables published by the Australian Bureau of Statistics¹⁰.

2.4 Second primary cancers

The volumes used to calculate the time-dependent second primary cancer induction probability (SPCIP) for each patient included the brain and the whole body (with the brain and tumour volumes subtracted). The SPCIP was then calculated for both of these volumes using the parameters determined by Schneider *et al.*¹¹. The yearly SPCIP derived for each volume was combined into a single probability for each year x after treatment as follows:

$$SPCIP(x) = 1 - (1 - SPCIP_{Body}(x))(1 - SPCIP_{Brain}(x)).$$

Treatment of second primary cancer was not considered. The yearly second cancer death probability was assumed to be 0.08, derived from 5-year survival rates of all cancers combined¹².

2.5 Injuries

Several injuries were considered in this analysis. These included brainstem necrosis, spinal cord myelitis, tinnitus (damage to the cochlea), blindness (damage in either optic nerve or

the optic chiasm), xerostomia (damage to the parotid glands), cataracts (damage to the lens), and endocrine dysfunction (damage to the pituitary gland). The model determined by Lee *et al.*¹³ was used to estimate the NTCP for tinnitus. The model determined by De Marzi *et al.*¹⁴ was used to estimate the NTCP for endocrine dysfunction. The models used for all other injuries have been described in previous work⁸.

It was assumed that all injuries were non-fatal, with the exception of brainstem necrosis, which was assumed to be fatal within one year for all patients affected.

All injuries were assumed to be chronic, with the exception of cataracts (which can usually be treated with surgery) and spinal cord myelitis due to a lack of data on long-term costs for this complication.

2.6 Estimation of costs and utilities

Costs associated with both the primary radiation treatment and treatment of side effects were incorporated into this model. Each state of health was assigned both a yearly cost and a quality of life utility. By default, the Well state has a utility of 1 and the Deceased state has a utility of 0. All other states have utilities within this interval depending on the impact of the corresponding complication on patient quality of life. The utilities used in this work are listed in Table 2, and are used to calculate the QALE. Second cancers were assumed to have a utility of 0.8. These assumptions are discussed in detail by Austin *et al.*⁸.

Table 2: Estimates for the quality of life utilities for states in the Markov model.

State	Utility	Comments
Base of skull chordoma	0.72 ¹⁵	
Second primary cancer	0.8	Not clinically founded
Brainstem necrosis	0.6	Not clinically founded
Spinal cord myelitis	0.7 ¹⁶	Utility for spinal cord stenosis taken as an approximation for myelitis
Tinnitus	0.58 ¹⁷	Evaluated after visiting a tinnitus clinic
Blindness	0.33 ¹⁶	Complete blindness
Xerostomia	0.83 ¹⁵	
Cataracts	0.6 ¹⁶	Advanced lens opacity
Endocrine dysfunction	0.73 ¹⁸	Utility of adult females with growth hormone deficiency. Average of values derived from Belgian and Dutch cohorts.

No costs were assumed for death, only loss of QALYs. Costs and utilities were discounted by 3% annually, to adjust for differences in timing of costs and effects. All costs are listed in Australian dollars. The costs applied in the model are as follows:

- **Radiation therapy:** The cost of a photon treatment was assumed to be \$11,877¹⁹ and the cost of PBT to be conservatively 2.5 times greater¹⁹.
- **Chordoma and second cancers:** Re-treatments are not incorporated into the model and hence the only assumed cost associated with cancers was due to lost productivity. The reduction in Australia's GDP has been found to be \$1,738 million due to 108,900 cancer patients not participating in the work force²⁰, or approximately \$15,960 per patient per year. The same estimate was applied for the chordoma state, due to a lack of data specific to this rare cancer. In this model, it is likely that the majority of patients in

cancer states will move to the Deceased state before they reach the typical retirement age or shortly after. To reduce bias however, the estimated cost of lost productivity was not applied to patients when they were older than 65 years, which is the minimum age to be eligible for the aged pension in Australia. In addition, costs due to lost productivity were not applied when the patient age was less than 18.

- **Spinal cord myelitis:** The cost of an episode of myelitis was assumed to be \$43,764²¹. This cost was only applied once, rather than annually.
- **Tinnitus:** The treatment and societal costs of tinnitus have been analysed by Maes *et al.*²² On average, the annual cost of tinnitus per patient was estimated as €11,949 (AUD\$18,918), representing a significant economic burden. Productivity losses were included in this estimate.
- **Blindness:** It has been estimated that vision impairment cost \$9.85 billion in Australia in 2004, corresponding to 480,000 vision-impaired people²³. This implies that on average, a vision-impaired person costs \$20,520 annually. This cost estimate includes both direct healthcare expenditure and indirect costs such as carer costs, lost earnings and welfare payments.
- **Xerostomia:** The annual cost of xerostomia was assumed to be US\$2,144²⁴ (AUD\$2,950), including oral saline rinses, pilocarpine, dental and nutritionist visits and fluoride gel.
- **Cataracts:** Cataracts was assumed to be treatable with surgery involving lens extraction and insertion of an intraocular lens. The cost of this was estimated at \$760 based on the Medicare Benefits Schedule²⁵.
- **Endocrine dysfunction:** The cost of the medicine required to treat GHD is on average \$5,478 per patient annually¹⁹. This cost was only applied to patients aged 18 years and

under as treatment is usually not necessary beyond this age. However, it was assumed that it was not possible to recover from this injury.

3 Results

The dose-dependent transition probabilities calculated for each plan and for each patient are summarised in Table 3. No patients had a significant risk of brainstem necrosis, spinal cord myelitis, or blindness from any treatment. This is most likely due to these organs being particularly critical and being weighted accordingly during plan optimisation. IMPT was able to provide a much greater probability of locoregional control in some patients (the greatest difference was 0.2 for Patient 1). Tinnitus and endocrine dysfunction were the most common injuries in the cohort, although the probabilities of these complications were negligible in some patients, regardless of the treatment. The risk of xerostomia was 10 times greater for Patient 2 if treated with VMAT (10% compared with 1%). Patient 1 had a 25% chance of developing cataracts if treated with VMAT compared with a negligible probability (<1%) if treated with IMPT.

The Markov model took the dose-dependent transition probabilities as input to calculate an ICER for each patient (Table 4). In accordance with NICE guidelines²⁶, an IMPT treatment was classified as cost-effective if it could be provided at a cost of £20,000-30,000 (AUD\$36,000-54,000) per QALY gained or less compared with VMAT. Table 4 demonstrates that all patients could be treated with PBT cost-effectively. The mean ICER was AUD\$3,220 per QALY gained. Of particular interest was Patient 5, who had an improved predicted clinical outcome if treated with IMPT. However, this also corresponded to a lower cost

compared with VMAT when complication costs were considered. This is likely a result of the elevated dose received by the ear with VMAT.

One-way sensitivity analyses were conducted to test the sensitivity of the results to estimated model parameters. The results are summarised in Table 5. Only the parameters associated with cancers, tinnitus, and endocrine dysfunction were considered in the sensitivity analysis, as these were the most common injuries. As the most likely driver of the treatment cost ratio is the PBT cost, this cost was varied in the sensitivity analysis. The fraction of the cohort that could be treated cost-effectively remained stable with all parameter variations.

Table 3: Dose-dependent transition probabilities for each patient. The second primary cancer induction probabilities (SPCIPs) listed represents the probability of a second primary cancer within 25 years after treatment. The normal tissue complication probabilities (NTCPs) represent the time integrated probabilities. TCP = tumour control probability.

Patient #	Treatment	TCP	Brainstem necrosis	Spinal cord myelitis	Tinnitus	Blindness	Xerostomia	Cataracts	Endocrine dysfunction	SPCIP
1	IMPT	0.88	<0.01	<0.01	0.98	0.01	<0.01	<0.01	1.00	0.01
	VMAT	0.68	<0.01	<0.01	0.98	<0.01	<0.01	0.24	1.00	0.02
2	IMPT	0.89	<0.01	<0.01	0.01	<0.01	0.01	<0.01	<0.01	0.01
	VMAT	0.82	<0.01	<0.01	0.01	<0.01	0.10	<0.01	<0.01	0.02
3	IMPT	0.94	0.02	0.02	0.01	<0.01	<0.01	<0.01	0.82	<0.01
	VMAT	0.75	<0.01	<0.01	0.13	<0.01	<0.01	<0.01	0.25	0.01
4	IMPT	0.86	0.01	<0.01	0.02	<0.01	<0.01	<0.01	0.39	0.01
	VMAT	0.78	<0.01	<0.01	0.04	<0.01	<0.01	<0.01	0.79	0.03
5	IMPT	0.83	0.01	0.01	0.48	<0.01	<0.01	<0.01	0.99	0.02
	VMAT	0.79	<0.01	<0.01	0.67	<0.01	<0.01	0.02	0.84	0.03
6	IMPT	0.62	<0.01	<0.01	0.01	<0.01	<0.01	<0.01	-	<0.01
	VMAT	0.61	<0.01	<0.01	0.03	0.01	<0.01	0.01	-	0.01
7	IMPT	0.83	<0.01	<0.01	0.05	<0.01	-	<0.01	0.97	0.01
	VMAT	0.74	<0.01	<0.01	0.14	0.01	-	0.01	0.83	0.02

Table 4: Predicted life expectancies, costs and ICERs for each patient. QALE = quality adjusted life expectancy; QALY = quality adjusted life year.

Patient ID	Treatment	Raw LE (years)	QALE (QALYs)	Cost (\$)	ICER (\$/QALY)
1	IMPT	68.4	29.8	358,320	13,620
	VMAT	52.8	23.1	266,440	
2	IMPT	63.8	63.2	35,320	2,230
	VMAT	58.7	57.2	21,860	
3	IMPT	71.2	55.3	72,350	4,220
	VMAT	57.8	51.0	54,020	
4	IMPT	34.1	30.3	37,790	2,110
	VMAT	31.3	24.4	25,390	
5	IMPT	47.6	28.5	189,990	-15,800
	VMAT	45.3	26.0	229,800	
6	IMPT	22.9	22.6	39,390	14,250
	VMAT	22.3	21.7	26,270	
7	IMPT	66.2	47.9	98,590	1,910
	VMAT	57.4	41.9	87,100	

Table 5: The effect of model parameter variation on the percentage of the cohort that could be treated with IMPT cost-effectively. The treatment cost ratios were altered by altering the proton treatment cost. This cost is more likely to vary compared with the photon treatment cost.

Scenario	Percentage cost-effective
No parameter variation	100
Decreased proton/photon cost ratio to 1.5	100
Increased proton/photon cost ratio to 3.5	100
<i>Primary cancer state</i>	
Decreased cost to 75%	100
Increased cost to 125%	100
Decreased utility by 0.1	100
Increased utility by 0.1	100
<i>Second primary cancer state</i>	
Decreased cost to 75%	100
Increased cost to 125%	100
Decreased utility by 0.1	100
Increased utility by 0.1	100
<i>Tinnitus state</i>	
Decreased cost to 75%	100
Increased cost to 125%	100
Decreased utility by 0.1	100
Increased utility by 0.1	100
<i>Endocrine dysfunction state</i>	
Decreased cost to 75%	100
Increased cost to 125%	100
Decreased utility by 0.1	100
Increased utility by 0.1	100

4 Discussion

For all of the cases presented, it was found that the initial cost of the proton treatment was justifiable if the costs associated with the greater risk of radiation-induced toxicity arising from photon treatments are considered. This was predominantly due to reduced risks of tinnitus and endocrine dysfunction, as well as improved tumour control probabilities

associated with the IMPT treatments. It was predicted that the proton treatment for one of the patients was cost-saving, that is, the treatment of both the tumour and treatment side effects were both less expensive and resulted in an improved clinical outcome compared with VMAT.

The results presented here are consistent with those of Mailhot Vega *et al.*⁴, in that PBT has been found to be a cost-effective treatment in cases where critical structures can be spared (typically the pituitary and cochlea in this case). While Peeters *et al.*⁵ found a larger cost difference between proton and photon treatments for BOSCh (AUD\$26,070) compared with other indications, here proton treatments were found to be cost-effective for BOSCh patients with a mean cost difference of AUD\$17,200. The discrepancy between the results presented here and those of Peeters *et al.* is possibly due to different healthcare systems, as well as our inclusion of costs associated with additional treatment complications.

This work has limitations that should be considered when interpreting the results. The quality of life utility associated with the endocrine dysfunction state was derived from an adult population and may not be representative of the quality of life experienced by a paediatric patient, which may influence the ICER calculated for certain patients in the cohort considered here. Furthermore, it is possible that endocrine dysfunction could be associated with costs other than that of treating GHD, and the cost assumed in this work could be underestimated as a result. However, the results were stable with variations in the costs associated with this injury. The assumed cost of spinal cord myelitis could also be underestimated as it did not include treatment of additional complications associated with the condition. However, the NTCP calculated for this injury was <0.01 for most patients and

did not exceed 0.02 for any patient or treatment, so it is unlikely that this assumption impacted the results.

No costs were assumed for premature death. Due to the large TCP difference between IMPT and VMAT treatments for many patients, this assumption likely underestimates the costs associated with VMAT. There was also difficulty in sourcing accurate injury development times, resulting in a degree of uncertainty in the costs and QALYs.

Model validation is an important step in the process of developing individualised patient selection strategies²⁷. The estimated SPCIPs for several patients in the cohort were comparatively low considering observations of second malignancy incidences in all treatment sites. These have been found to be 7.5% for photons and 5.3% for protons (median follow up of 6.7 years)²⁸. The model used to calculate the SPCIP in this work is yet to be validated, and this is the likely reason for the discrepancy between this work and the observations.

5 Conclusion

Markov modelling provides a means for timely assessment of new technologies in radiation oncology. This concept has been applied in the current work on an individual patient dosimetry basis for the assessment of cost-effectiveness of PBT for BOSCh. The model suggested all patients could be treated cost-effectively with PBT when compared to VMAT. Sensitivity analyses demonstrated the robustness of these results. This form of assessment may prove useful in guiding public health system support for patients to receive PBT in Australia.

Acknowledgements

The authors wish to thank Peter Rhodes for the early development of the project. The first author acknowledges the support of an Australian Government Research Training Program Scholarship. The third author acknowledges the support of ACEMS (ARC Centre of Excellence for Mathematical and Statistical Frontiers).

References

1. Tasman Radiation Oncology Group (TROG). The assesment of new radiation oncology technologies and treatments (ANROTAT) Project Final Report. Newcastle2012.
2. Fletcher CD, Unni KK, Mertens F. *Pathology and genetics of tumours of soft tissue and bone*: IARC; 2002.
3. Amichetti M, Cianchetti M, Amelio D, Enrici RM, Minniti G. Proton therapy in chordoma of the base of the skull: a systematic review. *Neurosurgical review*. 2009;32:403.
4. Mailhot Vega R, Kim J, Hollander A, et al. Cost effectiveness of proton versus photon radiation therapy with respect to the risk of growth hormone deficiency in children. *Cancer*. 2015;121:1694-1702.
5. Peeters A, Grutters JP, Pijls-Johannesma M, et al. How costly is particle therapy? Cost analysis of external beam radiotherapy with carbon-ions, protons and photons. *Radiotherapy and Oncology*. 2010;95:45-53.
6. Jäkel O, Land B, Combs SE, Schulz-Ertner D, Debus J. On the cost-effectiveness of carbon ion radiation therapy for skull base chordoma. *Radiotherapy and Oncology* 2007;83:133-138.
7. Lundkvist J, Ekman M, Ericsson SR, Jönsson B, Glimelius B. Cost-effectiveness of proton radiation in the treatment of childhood medulloblastoma. *Cancer*. 2005;103:793-801.
8. Austin AM, Douglass MJ, Nguyen GT, Penfold SN. A radiobiological Markov simulation tool for aiding decision making in proton therapy referral. *Physica Medica*. 2017;44:72-82.
9. Fagundes MA, Hug EB, Liebsch NJ, Daly W, Efird J, Munzenrider JE. Radiation therapy for chordomas of the base of skull and cervical spine: patterns of failure and outcome after relapse. *International Journal of Radiation Oncology* Biology* Physics*. 1995;33:579-584.
10. ABS. Life Tables, States, Territories and Australia, 2013-2015. 2016.
11. Schneider U, Sumila M, Robotka J. Site-specific dose-response relationships for cancer induction from the combined Japanese A-bomb and Hodgkin cohorts for doses relevant to radiotherapy. *Theoretical Biology and Medical Modelling*. 2011;8:27.
12. AIHW. Cancer in Australia. *Cancer series no. 101*. 2017.
13. Lee TF, Yeh SA, Chao PJ, et al. Normal tissue complication probability modeling for cochlea constraints to avoid causing tinnitus after head-and-neck intensity-modulated radiation therapy. *Radiation Oncology*. 2015;10:194.
14. De Marzi L, Feuvret L, Boulé T, et al. Use of gEUD for predicting ear and pituitary gland damage following proton and photon radiation therapy. *The British Journal of Radiology*. 2015;88:20140413.
15. Ramaekers BL, Grutters JP, Pijls-Johannesma M, Lambin P, Joore MA, Langendijk JA. Protons in head-and-neck cancer: bridging the gap of evidence. *International Journal of Radiation Oncology* Biology* Physics*. 2013;85:1282-1288.
16. Tengs TO, Wallace A. One thousand health-related quality-of-life estimates. *Medical care*. 2000;583-637.

17. Refaie AE, Davis A, Kayan A, Baskill J, Lovell E, Owen V. A questionnaire study of the quality of life and quality of family life of individuals complaining of tinnitus pre-and postattendance at a tinnitus clinic. *International Journal of Audiology*. 2004;43:410-416.
18. Busschbach J, Wolffenbuttel B, Annemans L, Meerding W, Kołtowska-Häggström M. Deriving reference values and utilities for the QoL-AGHDA in adult GHD. *The European Journal of Health Economics*. 2011;12:243-252.
19. Kemp R, Ramiscal R, Fodero L, Reardon O, Lyon S, Scuteri J. Assesment of Proton Beam Therapy: Commonwealth of Australia, Canberra, ACT; 2017.
20. Bates N, Callander E, Lindsay D, Watt K. Labour force participation and the cost of lost productivity due to cancer in Australia. *BMC Public Health*. 2018;18:375.
21. Bennett G, Dealey C, Posnett J. The cost of pressure ulcers in the UK. *Age and Ageing*. 2004;33:230-235.
22. Maes IH, Cima RF, Vlaeyen JW, Anteunis LJ, Joore MA. Tinnitus: a cost study. *Ear and Hearing*. 2013;34:508-514.
23. Taylor H, Pezzullo M, Keeffe J. The economic impact and cost of visual impairment in Australia. *British Journal of Ophthalmology*. 2006;90:272-275.
24. Bonomi A, Palmer C, Ajax M, Peeples P, Jackson S. Cost of managing mucositis and xerostomia in head and neck cancer patinets undergoing chemoradiohearpy or radiation. *Value in Health*. 1999;2:197.
25. Medicare Benefits Scheme Item number 42702.
26. NICE. Methods for the development of NICE public health guidance (third edition) - incorporating health economics. 2012.
27. Langendijk JA, Lambin P, De Ruysscher D, Widder J, Bos M, Verheij M. Selection of patients for radiotherapy with protons aiming at reduction of side effects: the model-based approach. *Radiotherapy and Oncology*. 2013;107:267-273.
28. Chung CS, Yock TI, Nelson K, Xu Y, Keating NL, Tarbell NJ. Incidence of second malignancies among patients treated with proton versus photon radiation. *International Journal of Radiation Oncology* Biology* Physics*. 2013;87:46-52.

6.3 Discussion and conclusion

The publication presented in this chapter demonstrates that a consideration of treatment cost-effectiveness has been successfully added to the patient selection tool. This will increase the utility of the model in future applications.

The results suggest that base of skull chordoma is not only a standard indication for proton therapy, but also a cost-effective one. As proton therapy becomes an available treatment option in Australia, such evidence could be used as a means of justifying the reimbursement of providing the treatment to base of skull chordoma patients.

Chapter 7

Cost-effectiveness of Proton Therapy: Breast Cancer

The publication P4 forms the basis of this chapter:

Austin, A.M., Douglass, M.J.J., Nguyen, G.T., Cunningham, L., Le, H., Hu, Y. & Penfold, S.N.. Individualised selection of left-sided breast cancer patients for proton therapy based on cost-effectiveness. *International Journal of Particle Therapy*. 2019 (Submitted May 2019).

7.1 Introduction and motivation

The method of incorporating treatment cost-effectiveness into the patient selection tool has been presented in Chapter 6. The output was demonstrated with a small cohort of base of skull chordoma patients. This disease is relatively rare and is most common in paediatric patients. In contrast, breast cancer is more prevalent. However, unlike base of skull chordoma, breast cancer is not considered to be a standard indication for proton therapy.

While the costs of construction and operation of proton therapy treatment facilities is significant, they become increasingly viable if more patients are expected to benefit from

the treatment. Therefore, the cost-effectiveness of treating common indications with proton therapy warrants further investigation.

The aim of the publication that forms the basis of this chapter was to apply the method of patient selection for proton therapy that had been developed in this work to a non-standard indication. In this case, the indication considered was left-sided breast cancer. Patients with this diagnosis that receive radiation therapy have an elevated risk of complications related to the heart, lungs and the contralateral breast. The possibility of a reduced dose to these organs offered by proton therapy suggests that left-sided breast cancer patients could benefit from the treatment. Another key difference between this cohort and the cohort considered in Chapter 6 is that no patients received proton therapy. All treatment plans were created for the purposes of retrospective analysis as part of research.

This chapter details additional data and technical aspects that were not included in the publication, including probability calculations (Section 7.3) and the estimation of costs (Section 7.4). A discussion and conclusion are given in Section 7.7.

7.2 Statement of contribution

7.2.1 Conception

The idea to apply the Markov model to a cohort of breast cancer patients first conceptualised by Scott Penfold. The method by which to implement this was developed by Annabelle Austin.

7.2.2 Realisation

The writing of the code, analysis, and sourcing of parameters was performed by Annabelle Austin. The organs on the CT scans were contoured by Hien Le and Yvonne Hu. The radiotherapy treatment plans used in the analysis were created by Lisa Cunningham.

7.2.3 Documentation

This paper was written by Annabelle Austin. Editing was performed by all authors.

Statement of Authorship

Title of Paper	Individualised selection of left-sided breast cancer patients for proton therapy based on cost-effectiveness.
Publication Status	<input type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input checked="" type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Austin, A.M., Douglass, M.J.J., Nguyen, G.T., Cunningham, L., Le, H., Hu, Y. & Penfold, S.N.. Individualised selection of left-sided breast cancer patients for proton therapy based on cost-effectiveness. <i>International Journal of Particle Therapy</i> . 2019 (Submitted May 2019).

Principal Author

Name of Principal Author (Candidate)	Annabelle Austin
Contribution to the Paper	Developed methods and code, sourced input model parameters from literature performed simulations and analysis, wrote manuscript and acted as corresponding author.
Overall percentage (%)	85%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 27/05/19


Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Scott Penfold
Contribution to the Paper	Provided supervision, advice, and assisted in manuscript evaluation.
Signature	Date 22/05/19

Name of Co-Author	Michael Douglass
Contribution to the Paper	Provided supervision, advice, and assisted in manuscript evaluation.
Signature	Date 22/05/19

Name of Co-Author	Giang Nguyen		
Contribution to the Paper	Provided supervision, advice, and assisted in manuscript evaluation.		
Signature		Date	24/05/2019

Name of Co-Author	Lisa Cunningham		
Contribution to the Paper	Created treatment plans used in analysis and assisted in manuscript evaluation.		
Signature		Date	24/5/19

Name of Co-Author	Hien Le		
Contribution to the Paper	Performed organ contouring to facilitate creation of treatment plans used in analysis and assisted in manuscript evaluation.		
Signature		Date	24/5/19

Name of Co-Author	Yvonne Hu		
Contribution to the Paper	Performed organ contouring to facilitate creation of treatment plans used in analysis and assisted in manuscript evaluation.		
Signature		Date	24/5/19

Individualised selection of left-sided breast cancer patients for proton therapy based on cost-effectiveness

Running title: Patient selection for proton therapy

Authors:

Annabelle M. Austin¹

Michael J. J. Douglass^{1,2}

Giang T. Nguyen³

Lisa Cunningham⁴

Hien Le⁴

Yvonne Hu⁴

Scott N. Penfold^{1,2}

¹Department of Physics, University of Adelaide, Adelaide, SA, Australia

²Department of Medical Physics, Royal Adelaide Hospital, Adelaide, SA, Australia

³School of Mathematical Sciences, University of Adelaide, Adelaide, SA, Australia

⁴Department of Radiation Oncology, Royal Adelaide Hospital, Adelaide, SA, Australia

Corresponding author:

Annabelle M. Austin

Department of Physics

School of Physical Sciences

University of Adelaide

North Terrace

Adelaide, 5005

SA, Australia

Email: annabelle.austin@adelaide.edu.au

Phone: 08 8313 5996

Number of figures and tables: 4

Abstract

Introduction: The significantly greater cost of proton therapy compared with X-ray therapy is frequently justified by the expected reduction in normal tissue toxicity. This is often true for indications such as paediatric and skull base cancers. However, the benefit is less clear for other more common indications such as breast cancer. This is due to uncertainty regarding the effect of a reduced dose in the chest region on clinical outcome. The aim of this work is to demonstrate an individualised selection method for proton therapy of left-sided breast cancer patients based on cost-effectiveness of treatment.

Patients and Methods: 16 left-sided breast cancer patients had a treatment plan generated for the delivery of intensity modulated proton therapy (IMPT) and of intensity modulated photon therapy (IMRT) with the deep inspiration breath hold (DIBH) technique. The resulting dosimetric data was used to predict probabilities of tumour control and toxicities. These probabilities were used in a Markov model to predict costs and the number of quality adjusted life years expected as a result of each of the two treatments.

Results: IMPT was not cost-effective for the majority of patients, but was cost-effective where there was a greater risk reduction of second malignancies with IMPT.

Conclusion: The Markov model predicted that IMPT with DIBH can only be cost-effective for selected left-sided breast cancer patients where IMRT would result in a significantly greater dose to normal tissue. The presented model may serve as a means of evaluating the cost-effectiveness of IMPT on an individual patient basis.

Keywords

Proton therapy, cost-effectiveness, breast cancer, Markov model, decision aid, radiobiological models

1 Introduction

Comparative planning studies have suggested that proton therapy has the potential to increase sparing of critical structures in the treatment of breast cancers for certain patients^{1,2}. However, Weber *et al.*² noted that the issues of treatment cost and availability for a common disease could limit the routine clinical use of protons in the post-operative treatment of breast cancer.

While proton therapy is a more expensive treatment than conventional X-ray therapy, it may be justified when costs other than that of the initial treatment are considered over the lifetime of a patient. For some patients, savings may be made if they are treated with proton therapy, even if the initial cost is greater. Lundkvist *et al.*³ have investigated whether improved outcomes for breast cancer patients who receive proton therapy are sufficient to justify a greater treatment cost. They found the treatment to be cost-effective for patients who had a high risk of developing a cardiac complication as a result of the radiation. Mailhot Vega *et al.*⁴ have developed an approach of selecting breast cancer patients to receive proton therapy based on risk of radiation-induced cardiac toxicity and proton treatment cost-effectiveness. Proton therapy was found to be cost-effective for cases where a woman had a cardiac risk factor and would receive a mean heart dose of greater than 5 Gy if treated with photons.

Deep inspiration breath hold (DIBH) with X-rays is becoming increasingly common for the treatment of left-sided breast cancer⁵. This technique can increase the distance between the breast and the heart, reducing heart dose and thereby the risk of radiation induced heart complications. For patients capable of the breath hold technique, the reduction in risk of radiation induced toxicity may be negligible with proton therapy. Proton therapy still has the potential to reduce dose to the lung and contralateral breast compared with DIBH with

X-rays, however^{6,7}. These organs are particularly sensitive to radiation induced second primary cancers⁸. It is important these organs at risk are included in an analysis of cost-effectiveness of proton therapy compared to state-of-the-art X-ray therapy.

The objective of the current work was to assess cost-effectiveness of proton therapy for a cohort of 16 DIBH-capable early stage breast cancer patients. In addition to cardiac toxicity, pneumonitis and second primary cancer induction were included in a Markov chain cost-effective analysis comparing intensity modulated proton therapy (IMPT) with hybrid three dimensional conformal radiotherapy (3DCRT)/intensity modulated radiotherapy (IMRT) X-ray radiotherapy. The transition probabilities of the Markov model were based on radiobiological models of tumour control probability (TCP), normal tissue complication probability (NTCP) and second primary cancer induction probability (SPCIP). These probabilities were derived on an individual basis from their comparative proton/X-ray radiotherapy treatment plans. The model was used to predict the likely outcome after a given treatment for each of the patients in the cohort in terms of life expectancy and quality adjusted life expectancy (QALE). Costs of primary and subsequent treatment were also included to determine the cost per quality adjusted life year (QALY) gained for protons compared to X-rays, also known as the incremental cost-effectiveness ratio (ICER).

2 Patients and Methods

2.1 Patient cohort and treatment planning

The cohort of patients considered in this retrospective study consisted of 16 female left-sided breast cancer patients treated at the Royal Adelaide Hospital with X-ray radiotherapy. The median age was 56 years (range 36-74). 50% of diagnoses were invasive ductal carcinoma, but diagnoses also included invasive lobular carcinoma, papillary carcinoma, ductal carcinoma in situ, and apocrine carcinoma. A majority (68%) of patients were stage

T1 and N0 (87.5%). As patients with metastatic disease were excluded from the study, most patients were M0, with one patient Mx (unable to be assessed for distant metastases). All patients had breasts intact and the whole breast was modelled in treatment planning.

Each patient had a computed tomography (CT) scan acquired with DIBH. The clinical target volume (CTV) included apparent CT glandular breast tissue and lumpectomy CTV. In this retrospective analysis, each patient had two new treatment plans created. The prescribed dose was 40 Gy_{RBE} (the RBE weighted dose) in 15 fractions. Planning objectives for the heart were $D_{\text{mean}} < 3 \text{ Gy}$ and $V_{21.5\text{Gy}} < 10\%$, for the left lung $V_{18\text{Gy}} < 15\%$ and as low as reasonably achievable doses to the left anterior descending artery, right lung, and right breast.

The X-ray treatment plan made use of the 3DCRT/IMRT hybrid technique (h-IMRT). The plans consisted of opposing tangential fields with 70% and 30% weighting of the 3DCRT (open tangent field) and IMRT beams (inversely optimized IMRT field) in each tangent respectively. This weighting was used to ensure planning consistency. 6 MV beams were used unless the size of the breast required the use of 10 MV beams in the 3DCRT beam to reduce lateral hotspots and improve target coverage. IMRT beams were optimized to a planning target volume, defined as the CTV with a 5 mm margin limited to within the exterior of the patient minus 5 mm and excluding the left lung. Treatment plans were created to achieve 98% coverage of the planning target volume (PTV) with 95% of the prescribed dose.

IMPT plans were created with a single *en-face* beam. A range-shifter was used to allow the placement of Bragg peaks close to the patient surface. A beam specific PTV was generated with an expansion of 5 mm laterally and 3% of the beam range distally. Two patients were duplicated and re-planned to test planning consistency.

2.2 Markov model

A discrete-time Markov chain model was applied in this work⁹. The time period modelled begins immediately after the final fraction of radiotherapy treatment and ends when the patient is deceased. The cycle length was chosen to be one year.

A patient can occupy only a single Markov state at a given time. These states are summarised in Figure 1. The toxicities considered included pneumonitis and heart disease. The possibility of developing a second primary cancer (SPC) as a result of the initial radiation treatment was also included.

The following assumptions were made when determining the Markov states to be used in the model:

- Pneumonitis, if developed, is likely to resolve many years prior to the induction of a second primary cancer. Therefore there are no states for the situations where a patient is affected by both pneumonitis and a second primary cancer.
- If a treatment is unsuccessful, it is assumed that it is highly unlikely that the patient will still be alive when the probability of developing a second malignancy is significant. Therefore there are no states where the initial cancer and a second cancer coexist.

2.3 Markov state transition probabilities

Markov models in medical applications assume that a patient occupies a single state for the duration of a cycle. At the end of each cycle, it is possible for a patient to transition to another state. The allowed transitions are summarised in Figure 1.

The following assumptions were made when determining the allowed transitions in the model:

- It is not possible to recover from heart disease or a second cancer once it has developed.

- It is not possible to transition from the Cancer state to the Well state once the first Markov cycle has begun. The patient may begin in the Well state as a result of a successful treatment.
- There is a large difference in the time point after treatment at which the second primary cancer induction probability (SPCIP) becomes significant compared to the time point where the NTCP is significant for the toxicities considered. Therefore, the probability of simultaneously developing an injury and second cancer is negligible. Similarly, the probability of developing an injury after a second cancer is also negligible.
- Once pneumonitis has been recovered from, the model ensures that it is not possible to relapse.

The transition probabilities are explained in more detail in Sections 2.3.1 – 2.3.4.

[2.3.1 Locoregional control](#)

The probability of the patient beginning in the Well state is equal to the dose-dependent TCP (defined in the Supplementary Material), while the probability of beginning in the Cancer state is the complement of the TCP. Re-treatments are not directly included in the model. This is assumed as the outcome of the initial treatment is the focus of this selection tool. Similarly, while there were no explicit states for metastases, the cancer death probabilities incorporate this implicitly in the model.

[2.3.2 Normal tissue complications](#)

The probability of pneumonitis is calculated using the radiation dose to both lungs. The probability of heart disease is calculated with the dose to the heart. These transition probabilities are time-dependent to allow for a more realistic estimation of costs and QALYs. The details of the calculations are described in the Supplementary Material.

The majority of patients with pneumonitis recover¹⁰. It was assumed that recovery would occur after 1 year as all estimated costs associated with treating this injury applied within the first year only. It was assumed that heart disease was chronic and the possibility of recovery was neglected in this work.

2.3.3 Second primary cancer induction

The SPCIP is the probability of developing a radiation induced cancer as a result of the treatment. This is an important consideration due to the expected difference in integral dose between a proton and photon plan. This is also a dose- and time-dependent quantity and was calculated using the model developed by Schneider *et al.*¹¹. The relevant formula and input data are described in the Supplementary Material.

2.3.4 Death Probabilities

Unlike the other transition probabilities described in this section, the probability of transitioning to the Deceased state is dose-independent. The purpose of this study was to evaluate the cost-effectiveness from a dosimetric point of view and while the death probabilities are not dosimetric quantities, they allow for a more realistic estimate of the number of QALYs gained as a result of a given treatment. Depending on the Markov state of a patient, there are a number of possible transitions that can be made to the Deceased state:

- Death due to breast cancer as a result of an unsuccessful treatment. Survival of breast cancer patients was found to be 55% at 10 years for the case of local failure¹². A constant yearly death probability, denoted $Pr(Die)$, was derived from this data, where $S = (1 - Pr(Die))^n$, $n = 10$ is the number of years after treatment, and $S=0.55$ is the surviving fraction. Solving for the death probability gives 0.06.

- Death due to a second malignancy. The 5-year survival for all cancers combined is 68%¹³. Using the same method for the breast cancer death described above, a yearly death probability of 0.08 was derived.
- Death due to heart disease. The probability was assumed to be 0.01 per year which was estimated using 2007 prevalence (3.5 million) and death rates (48, 456) associated with cardiovascular disease in Australia¹⁴.
- Unrelated death. This time-dependent probability is based on data from life tables obtained from the Australian Bureau of Statistics (ABS)¹⁵.

Note that as recovery from pneumonitis is highly likely, it was assumed that this injury was non-fatal.

2.4 Estimation of quality of life utilities

The quality of life (QoL) utility value of each Markov state represents the quality of life associated with the state relative to perfect health (with QoL=1). By default, the quality of life associated with death is 0. The utilities used in the current work are listed in Table 1.

For states where there is more than one injury or cancer, the assigned utility is a multiplication of the utilities of the states where there is only one of each injury or cancer. The state representing the cases of second primary cancers were assigned a value of 0.8 in accordance with the Eastern Cooperative Oncology Group (ECOG) performance status¹⁶ as their definition of a grade 1 complication gives the most accurate description of this state.

2.5 Estimation of costs

In addition to the cost of the breast cancer treatment, costs of side-effect treatments were also incorporated into the model to allow for a more realistic representation of the costs associated with a given treatment. The costs of re-treatments and treatments of second

cancers were not included. No costs were assumed for fatal events, only loss of QALYs. For states where several injuries or cancers affect a patient, the cost applied was the sum of the costs for the individual injuries. Costs and QALYs were discounted by 3% annually, to adjust for differences in timing of costs and effects. Where possible, Australian costs were applied for consistency and all costs are in Australian dollars (AUD). These are summarised in Table 1. The details of the cost estimation are given in the Supplementary Material.

3 Results

The ICER was calculated for each patient in the cohort. The results are given in Table 2. In accordance with the NICE guidelines¹⁷, IMPT was considered cost-effective if it cost £20,000 (\$36,000) per QALY gained or less compared with h-IMRT.

Proton therapy was cost-effective for one patient in the cohort and cost-ineffective for 15 patients. Both members of both sets of the duplicated patients were classified as cost-ineffective. The difference in the ICER calculated for patient 3 and the ICER calculated for its re-planned duplicate was approximately \$8,000. For patient 8, the difference between the ICER and the ICER of the duplicate was \$6,000. These differences are due to small differences (up to 0.1 years) in the number of QALYs in the denominators.

3.1 Sensitivity analysis

Variation of selected parameters altered the fraction of the cohort that could be treated with IMPT cost-effectively.

Parameters related to second cancers were varied as these had a large impact on whether a patient was classified as cost-effective. These included costs, the utility and the death probability. In contrast, the TCP difference between the treatments did not exceed 0.01 for

any of the patients. Therefore, it is unlikely that variation of related parameters would impact the results.

Due to the relatively small NTCP for pneumonitis (see Supplementary Material), no parameters relevant to pneumonitis were considered to have a significant effect on the results. Even if the NTCP difference were larger between the two treatments, the relatively small cost and duration of pneumonitis would result in a minimal effect on the results. The exception was the possibility of this injury becoming chronic in a fraction of patients.

There was not a significant difference in the heart disease NTCP between IMPT and h-IMRT for any of the patients (see Supplementary Material). The only parameter related to heart disease that was varied was the baseline risk. This parameter was doubled in the analysis to investigate whether high risk groups could be treated with protons cost-effectively. Treatment cost ratios were varied by varying the proton treatment cost, as this was considered to have the greatest uncertainty.

After selecting parameters that were most likely to influence the results, a sensitivity analysis was performed for each. The results are presented in Table 3. As expected, if IMPT could be delivered at a lower cost (1.5 times that of IMRT), then a significantly greater proportion of the cohort could be treated with IMPT cost-effectively. Proton therapy was also less likely to be cost-effective where there was a reduced probability of death due to second cancer induction. The results were stable with variation of other model parameters.

4 Discussion

The Markov model predicted that IMPT could not be delivered cost-effectively to the majority of patients in the cohort investigated. The patient that could be treated cost-

effectively had a comparatively high lung dose (see Supplementary material) which increased the second cancer risk. The higher lung dose was necessary to spare breast tissue in this patient who had relatively larger breasts. This was also a younger patient (less than the assumed retirement age at the time of treatment) and hence in the model they had the potential to be less productive in society as a result of second malignancies. Alternatively, the difference in normal tissue doses between treatments was smaller for the remainder of the cohort.

The sensitivity analysis indicated that, as expected, the initial cost of the proton treatment had the largest impact on whether a patient could be treated cost-effectively. However, it is anticipated that the cost of proton therapy will decrease over time as it is a newer treatment. Furthermore, it is likely that the initial cost of building a proton clinic would have a large contribution to this cost. This cost can be increasingly justifiable with an increasing number of patients who are expected to benefit from the treatment. If breast cancer patients could be included in this category, then proton clinics may be more viable as current standard indications are predominantly relatively rare or paediatric cases.

Lung cancer was found to be the most likely second cancer in this work, agreeing with a study of second cancer incidences after X-ray therapy for breast cancer¹⁸. In a planning study of various X-ray treatment techniques for breast cancer, Santos *et al.*⁸ also found the lungs to have the highest second cancer risk.

The mean heart dose did not exceed 4 Gy for any treatment for any of the patients (see Supplementary material). This is likely a result of the DIBH technique, which is designed to reduce exposure to the heart. Mailhot Vega *et al.*⁴ found that for a proton treatment of breast cancer to be cost-effective, it was necessary for the mean dose to the heart from

photons to be greater than 5 Gy. Hence, the results presented here are consistent with this finding.

The average predicted ICER of \$84,600 was smaller than the average predicted by Lundkvist *et al.* of €67,000 (\$105,000)³. Our estimation of the ratio of proton therapy to photon therapy costs is similar (2.5 in this work compared with 2.6). Their estimation of the probability of death due to breast cancer was lower than ours, but it is unlikely that this alone would influence our results significantly due to the relatively small difference in the expected TCP between IMPT and IMRT for the patients in our cohort. Therefore, the discrepancy is likely due to our inclusion of costs associated with the possibility of radiation-induced cancers.

While an ICER of £20,000 was assumed to be the threshold for a treatment to be cost-effective in this work, according to the NICE guidelines¹⁷ the threshold can be as large as £30,000 (\$54,000) if advisory bodies can make a strong case in support of the intervention. If this threshold were to be assumed here, an additional 4 patients would have a cost-effective proton treatment (31% of the cohort in total). These patients had relatively large lung dose differences between the two modalities, corresponding to larger SPCIP differences.

There are several assumptions in the Markov model that may have influenced the results. For example, re-treatments were omitted as the alternate treatments of the initial cancer are the subject of comparison in the model. However, the results may be less realistic as a consequence of this assumption. In reality re-treatments would likely occur and this would contribute to costs. In addition, loss of life is assumed to have no cost. Including each of

these factors would increase the likelihood of a proton treatment being cost-effective, assuming it resulted in improved tumour control and reduced second cancer rates.

The radiobiological models that are built into the Markov model also have limitations. For example, the model used to estimate the probability of developing heart disease was developed using data from both left and right-sided breast cancer patients. The effect of this is that the true NTCP may be underestimated, which could have contributed to the relatively small probabilities that were obtained for each of the patients despite a wide variation in mean heart dose.

It is worth noting that the patients considered in this study represent a subset of breast cancer patients who are able to hold their breath during treatment. This is not the case for all breast cancer patients, particularly those who are elderly. It may be possible to treat patients cost-effectively if they are not able to hold their breath or have suspected nodal involvement and therefore would experience a higher risk of cardiac toxicity if treated with X-rays.

5 Conclusion

The cost-effectiveness of proton therapy for a cohort of left-sided breast cancer patients capable of being treated with DIBH has been assessed with a Markov model. It was found that proton therapy was not a cost-effective treatment for the majority of the cohort. However, patients that would have an elevated risk of second malignancy if treated with X-rays could be treated with IMPT cost-effectively. The presented model has the potential to evaluate the cost-effectiveness of treatments on a case-by-case basis, facilitating the delivery of individualised medicine and ensuring the efficient usage of healthcare resources.

Acknowledgements

The authors wish to thank Peter Rhodes for the early development of the project. The first author acknowledges the support of an Australian Government Research Training Program Scholarship. The third author acknowledges the support of ACEMS (ARC Centre of Excellence for Mathematical and Statistical Frontiers).

References

1. Lomax AJ, Cella L, Weber D, Kurtz JM, Miralbell R. Potential role of intensity-modulated photons and protons in the treatment of the breast and regional nodes. *International Journal of Radiation Oncology* Biology* Physics*. 2003;55:785-792.
2. Weber DC, Ares C, Lomax AJ, Kurtz JM. Radiation therapy planning with photons and protons for early and advanced breast cancer: an overview. *Radiation Oncology*. 2006;1:22.
3. Lundkvist J, Ekman M, Ericsson SR, Isacson U, Jönsson B, Glimelius B. Economic evaluation of proton radiation therapy in the treatment of breast cancer. *Radiotherapy and Oncology*. 2005;75:179-185.
4. Mailhot Vega RB, Ishaq O, Raldow A, et al. Establishing cost-effective allocation of proton therapy for breast irradiation. *International Journal of Radiation Oncology* Biology* Physics*. 2016;95:11-18.
5. Bruzzaniti V, Abate A, Pinnarò P, et al. Dosimetric and clinical advantages of deep inspiration breath-hold (DIBH) during radiotherapy of breast cancer. *Journal of Experimental & Clinical Cancer Research*. 2013;32:88.
6. Lin LL, Vennarini S, Dimofte A, et al. Proton beam versus photon beam dose to the heart and left anterior descending artery for left-sided breast cancer. *Acta Oncologica*. 2015;54:1032-1039.
7. Cunningham L. *A retrospective dosimetric comparison of proton and X-ray therapy with deep inspiration breath hold for left-sided early stage breast cancer: a pilot study.*: School of Health Sciences, University of South Australia; 2018.
8. Santos AM, Marcu LG, Wong CM, Bezak E. Risk estimation of second primary cancers after breast radiotherapy. *Acta Oncologica*. 2016;55:1331-1337.
9. Austin AM, Douglass MJ, Nguyen GT, Penfold SN. A radiobiological Markov simulation tool for aiding decision making in proton therapy referral. *Physica Medica*. 2017;44:72-82.
10. Ghafoori P, Marks LB, Vujaskovic Z, Kelsey CR. Radiation-induced lung injury. Assessment, management, and prevention. *Oncology (Williston Park)*. 2008;22.
11. Schneider U, Sumila M, Robotka J. Site-specific dose-response relationships for cancer induction from the combined Japanese A-bomb and Hodgkin cohorts for doses relevant to radiotherapy. *Theoretical Biology and Medical Modelling*. 2011;8:27.
12. Fortin A, Larochelle M, Laverdière J, Lavertu S, Tremblay D. Local failure is responsible for the decrease in survival for patients with breast cancer treated with conservative surgery and postoperative radiotherapy. *Journal of Clinical Oncology*. 1999;17:101-101.
13. AIHW. Cancer in Australia. *Cancer series no. 101*. 2017.
14. AIHW. Cardiovascular disease: Australian facts. *Cardiovascular disease series*. 2011;53.
15. ABS. Life Tables, States, Territories and Australia, 2013-2015. 2016.
16. Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *American journal of clinical oncology*. 1982;5:649-655.
17. NICE. Methods for the development of NICE public health guidance (third edition) - incorporating health economics. 2012.
18. Grantzau T, Mellekjær L, Overgaard J. Second primary cancers after adjuvant radiotherapy in early breast cancer patients: a national population based study under the Danish Breast Cancer Cooperative Group (DBCG). *Radiotherapy and Oncology*. 2013;106:42-49.

19. Grann VR, Panageas KS, Whang W, Antman KH, Neugut AI. Decision analysis of prophylactic mastectomy and oophorectomy in BRCA1-positive or BRCA2-positive patients. *Journal of Clinical Oncology*. 1998;16:979-985.
20. Sullivan PW, Ghushchyan V. Preference-based EQ-5D index scores for chronic conditions in the United States. *Medical Decision Making*. 2006;26:410-420.

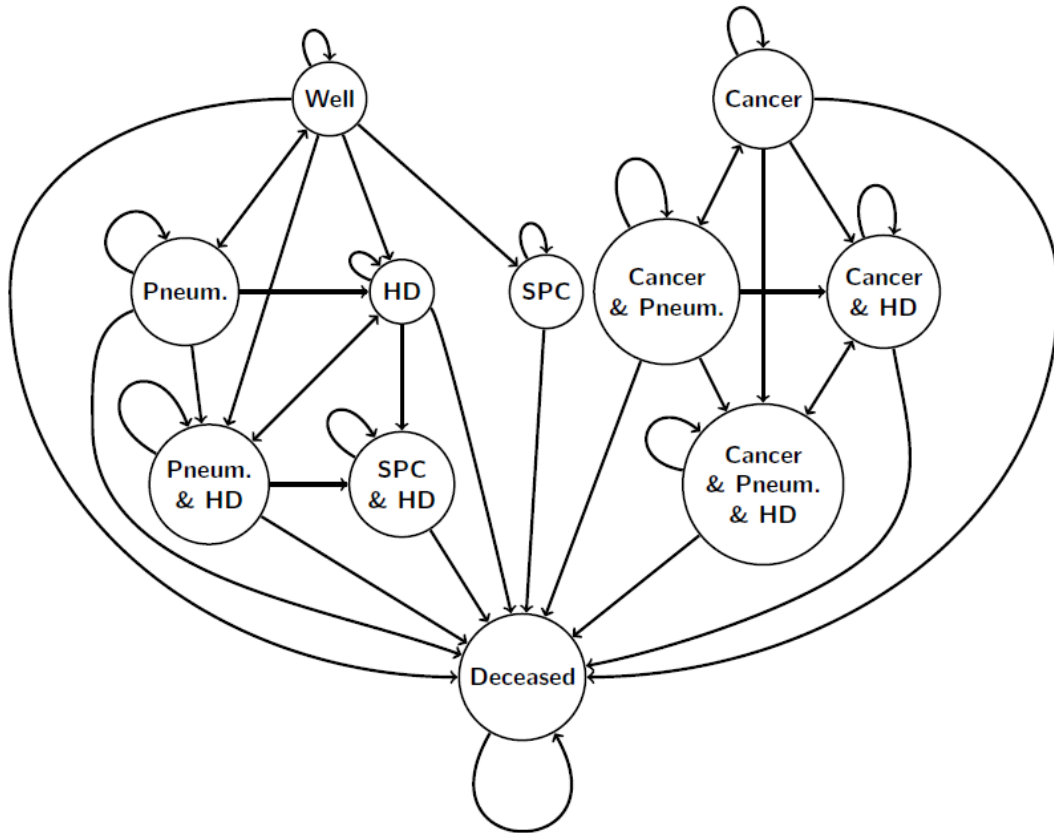


Figure 1: The Markov state transition diagram showing the allowed transitions between states. 'Well' represents perfect health. 'Cancer' represents the situation where the patient still has the initial primary breast cancer, 'SPC' represents a second primary cancer, 'Pneum.' refers to pneumonitis and 'HD' represent heart disease.

Table 1: Estimates for the yearly costs and quality of life utilities for states in the Markov model. Details of the cost estimations are given in the supplementary material.

State	Utility	Cost (\$)
Breast cancer	0.89 ¹⁹	15,960
Heart disease	0.8 ²⁰	13,658
Pneumonitis	0.8 ²⁰	4,037
Second primary cancer	0.8 ¹⁶	15,960

Table 2: Predicted life expectancies and costs for each patient. The cost of protons per QALY gained is listed. Duplicated patients are denoted by an asterisk. QALY = quality adjusted life year.

QALE = quality adjusted life expectancy=number of QALYs lived.

Patient ID	Treatment	Raw LE (y)	QALE (QALYs)	Costs (\$)	ICER (\$/y)
1	IMPT	48.2	25.48	39,040	73,950
	IMRT	47.6	25.26	22,470	
2	IMPT	19.0	14.38	31,430	107,130
	IMRT	18.8	14.21	13,650	
3	IMPT	38.9	22.91	38,040	122,700
	IMRT	38.6	22.77	20,740	
3*	IMPT	38.9	22.88	38,160	128,640
	IMRT	38.6	22.75	21,050	
4	IMPT	22.3	16.16	31,720	49,610
	IMRT	21.9	15.80	13,960	
5	IMPT	19.0	14.40	31,430	79,660
	IMRT	18.8	14.17	13,660	
6	IMPT	33.4	21.02	37,000	89,980
	IMRT	33.1	20.83	20,000	
7	IMPT	25.6	17.78	33,760	46,640
	IMRT	25.2	17.41	16,450	
8	IMPT	33.5	21.07	36,490	54,730
	IMRT	33.0	20.77	20,180	
8*	IMPT	33.4	21.01	37,050	60,910
	IMRT	33.0	20.74	20,550	
9	IMPT	18.2	13.94	31,360	67,800
	IMRT	18.0	13.67	13,600	
10	IMPT	21.4	15.69	31,650	87,370
	IMRT	21.2	15.48	13,910	
11	IMPT	36.2	22.00	37,660	237,110
	IMRT	36.0	21.92	20,110	
12	IMPT	15.0	12.03	31,040	99,620
	IMRT	14.9	11.85	13,410	
13	IMPT	18.1	13.92	31,360	90,840
	IMRT	18.0	13.73	13,920	
14	IMPT	30.7	19.96	36,450	26,750
	IMRT	29.9	19.38	20,780	
15	IMPT	40.6	23.42	38,660	52,680
	IMRT	40.0	23.12	22,540	
16	IMPT	29.0	19.26	35,480	45,820
	IMRT	28.5	18.90	18,800	

Table 3: The effect of model parameter variation on the percentage of the cohort that could be treated with IMPT cost-effectively. The ratio of the treatment costs was varied by varying the proton treatment cost.

Scenario	Percentage cost-effective
No parameter variation	6
Decreased proton/photon cost ratio to 1.5	88
Increased proton/photon cost ratio to 3.5	0
Decreased recovery rate of pneumonitis to 80%	6
Decreased second cancer death probability by 50%	0
Increased second cancer death probability by 50%	6
Decreased second cancer cost to 75%	6
Increased second cancer cost to 125%	6
Decreased second cancer utility by 10%	6
Increased second cancer utility by 10%	6
Doubled baseline risk of heart disease	6

7.3 Calculation of transition probabilities from DVH data

7.3.1 Tumour control probability (TCP) calculation

The TCP used in this model is based on the linear quadratic (LQ) model [58]:

$$\text{TCP} = \prod_{i=1}^M P(D_i)^{v_i}, \quad (7.1)$$

$$\text{with } P(D_i) = \exp\left(-\exp\left(e\gamma - \alpha D_i - \beta \frac{D_i^2}{n_{frac}}\right)\right), \quad (7.2)$$

where there are a total of M voxels, each having a fractional volume v_i (of the total tumour) that receives dose D_i as part of a treatment delivered in n fractions (n_{frac}), and α and β are the linear and quadratic coefficients of the LQ model, respectively. For breast cancer, $\alpha/\beta = 2.88$ and $\alpha = 0.08$ [91]. The dose was converted to an equivalent dose in 2 Gy fractions.

The parameter γ is the normalised dose-response gradient evaluated at $D = D_{50}$, the dose at which 50% of tumours are controlled. This value was taken as 1.46 [92].

7.3.2 Normal tissue complication probabilities

7.3.2.1 Pneumonitis

The Lyman-Kutcher-Burman (LKB) NTCP formalism was used to determine the probability of developing pneumonitis. The NTCP is given by (7.3) and (7.4) [60, 61],

$$\text{NTCP (pneumonitis)} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t \exp\left(\frac{-x^2}{2}\right) dx, \quad (7.3)$$

$$\text{with } t = \frac{D_{eff} - TD_{50}}{mTD_{50}} \quad \text{and} \quad D_{eff} = \left(\sum_{i=1}^{\ell} v_i D_i^{\frac{1}{n}}\right)^n, \quad (7.4)$$

where TD_{50} is the uniform dose in Gy given to the entire organ that results in 50% complication risk, m is an organ-specific parameter that represents the gradient of the dose-response curve (analogous to γ in the TCP calculation), and n is a parameter that characterises the volume dependence of the organ's response to radiation, and ℓ is the number of voxels. The values determined by Seppenwoolde *et al.* [93] ($TD_{50} = 30.8$ Gy, $m = 0.37$, $n = 0.99$) were used in this model. An α/β of 3.0 was assumed for the lung to convert the dose to an equivalent dose in 2 Gy fractions.

To calculate a time-dependent NTCP, a time-dependent normal distribution was defined that was normalised such that the integral was equal to the NTCP defined in (7.3). The mean of the distribution was the mean time taken for pneumonitis to develop. This was assumed to be 6 months with a standard deviation of 2 months as most cases are expected to develop within a year [94]. The result is a discretised normal distribution with a NTCP for each year after treatment.

7.3.2.2 Heart disease

The probability of a major coronary event was calculated using the model developed by Darby *et al.* [27]:

$$\text{NTCP (heart disease)} = B(1 + KD), \quad (7.5)$$

where $K = 0.074$ Gy, D is the mean dose to the heart, calculated from the individual patient's DVH for the heart associated with a particular plan, and B is the risk of a cardiac event without radiation therapy. This was estimated to be 1.5% based on a review of the prevalence of heart failure in Australia [95], which was found to be 1-2%. An α/β of 3.0 was assumed for the heart to convert the dose to an equivalent dose in 2 Gy fractions.

Once the NTCP for cardiac events was calculated, time-dependent probabilities were derived using the same method described for pneumonitis. The mean time taken for a cardiac event to occur after treatment was estimated to be based on data presented by Darby *et al.*[27] on the percentage increase of events per Gy for given time periods after radiotherapy. The overall average of the mean dose to the heart in their population study

was 4.9 Gy. The largest increase was within 4 years of radiotherapy, and hence the mean was assumed to be 4 years. A standard deviation of 2 year was assumed in this model, but this was not clinically founded.

7.3.3 Second primary cancer induction probabilities

The excess absolute risk (EAR) of developing a cancer in a particular organ at a particular time after treatment due to radiation exposure is given by (7.6) and was taken as an estimate of the SPCIP [84, 96]

$$EAR^{org}(age) = \frac{1}{V_T} \sum_{i=1}^{\ell} v_i(D_i) \beta_{EAR} RED(D_i) \mu(ageX, age). \quad (7.6)$$

Here, $ageX$ is the age of the patient at the time of treatment (the time of exposure to radiation), age is the age of the patient after treatment at the year of interest, V_T is the total volume of the organ, β_{EAR} is the initial slope, ℓ is the number of voxels, and $\mu(ageX, age)$ is the modifying function,

$$\mu(ageX, age) = \exp \left[\gamma_e (ageX - 30) + \gamma_a \ln \left(\frac{age}{30} \right) \right], \quad (7.7)$$

with γ_e and γ_a being the age modifying parameters.

Equation (7.8) gives the risk equivalent dose (RED) mechanistic model which accounts for the effects of cell killing and fractionation

$$RED(D) = \frac{e^{-\alpha'D}}{\alpha'R} \left(1 - 2R + R^2 e^{\alpha'D} - (1 - R)^2 e^{\frac{\alpha'R}{1-R}D} \right), \quad (7.8)$$

where R is the repopulation/repair parameter, α' is given by

$$\alpha' = \alpha + \beta \frac{D_i}{D_T} d_T, \quad (7.9)$$

and D_T and d_T represent the prescribed dose to the target volume and the corresponding dose per fraction, respectively. The values used for these parameters are listed in Table 7.1.

It was assumed that it was possible to develop a second cancer in any of the following: the left lung, the right breast, and all regions in the body that were not contoured. The parameters used for the SPCIP calculation are listed in Table 7.1 [84].

TABLE 7.1: Parameters used for the calculation of the SPCIP for each year after treatment. Data was not provided for the heart.

SPCIP PARAMETERS					
Tissue	β_{EAR}	α	R	γ_e	γ_a
Lung	8.0	0.042	0.83	0.002	4.23
Body	74.0	0.089	0.17	-0.024	2.38
Breast	8.2	0.044	0.15	-0.037	1.7

Once the SPCIP for each year had been calculated for the two tissues being considered, they were combined into a single probability of developing any second cancer (SC) in a given year, $P(\text{SC } 1 \cup \text{SC } 2)$, assuming the two events were independent.

7.4 Estimation of costs

7.4.1 Radiation therapy

The cost of IMRT was assumed to be \$11,877 [14] and the cost of proton therapy to be conservatively 2.5 times greater [14].

7.4.2 Cancer

The estimate of the cost associated with an unsuccessful treatment (resulting in the patient being in a state associated with cancer) was based on lost productivity (see Section 2.6 of the publication P3, Chapter 6). The model does not incorporate the effect of re-treatments and hence the costs associated with this were not included. In reality, it is likely that re-treatments would occur resulting in the true cost being higher. However, the original treatment (either proton or photon) is the subject of comparison in this study.

7.4.3 Heart disease

In the 2004-2005 financial year, \$5,942 million was spent on approximately 685,000 cardiovascular disease patients in Australia, corresponding to an average cost per year per person of \$8,674. This estimate includes hospital admissions, prescription pharmaceuticals (including the contribution from the PBS (pharmaceutical benefits scheme) and patient contributions), and out-of-hospital medical services [97]. The expenditure on lipid-lowering medicines was not included which may lead to a significant underestimation of costs. The annual expenditure on research (\$164 million) was subtracted for this study as it is not necessarily case specific.

To estimate costs associated with lowered productivity and workforce participation, the cost of heart disease-related absenteeism in 2004 obtained for a study in productivity loss [98] was applied. Applying a prevalence of 355,600 in 2004 [99], the total cost of \$31.7 million was translated to \$89 per person. Similarly to cancer, costs associated with workforce participation were applied only when the patients were below the retirement age.

7.4.4 Pneumonitis

The costs associated with pneumonitis were estimated based on management strategies described by Ghafoori *et al.* [100]. These included a chest X-ray and a course of oral prednisone (a corticosteroid). The cost of the X-ray was estimated to be \$47, based on

Medicare Benefits Schedule (MBS) data [101]. The medication cost was estimated at \$63, based on the Pharmaceutical Benefits Scheme (PBS) dispensed price [102].

In the estimations of Lundkvist *et al.* [40], one month of leave was assumed for 75% of patients with the injury. This corresponded to a total average leave cost of \$3,927 after applying data relevant to Australia. A patient recovers from pneumonitis after a single cycle and hence the cost is applied for one year only.

7.5 Dose-dependent transition probabilities by patient

The transition probabilities that were calculated based on the dose received by each patient are summarised in Table 7.2. The NTCPs and SPCIP are time-dependent probabilities. The total NTCP integrated from the starting age of the patient to the maximum possible age in the model (100) and the average yearly SPCIP is listed.

TRANSITION PROBABILITIES

Patient ID	Treatment	TCP	Pneumonitis NTCP	Heart NTCP	SPCIP
1	IMPT	0.971	<0.01	0.02	0.053
	IMRT	0.973	0.01	0.02	0.130
2	IMPT	0.971	<0.01	0.02	0.044
	IMRT	0.974	0.01	0.02	0.110
3	IMPT	0.971	<0.01	0.02	0.055
	IMRT	0.975	0.01	0.02	0.110
3 *	IMPT	0.972	<0.01	0.02	0.065
	IMRT	0.972	0.01	0.02	0.100
4	IMPT	0.971	<0.01	0.02	0.044
	IMRT	0.975	0.02	0.02	0.160
5	IMPT	0.971	<0.01	0.02	0.040
	IMRT	0.971	0.01	0.02	0.110
6	IMPT	0.973	<0.01	0.02	0.069

	IMRT	0.971	0.01	0.02	0.120
7	IMPT	0.971	<0.01	0.02	0.057
	IMRT	0.973	0.02	0.02	0.170
8	IMPT	0.976	<0.01	0.02	0.067
	IMRT	0.972	0.01	0.02	0.130
8 *	IMPT	0.971	<0.01	0.02	0.069
	IMRT	0.969	0.01	0.02	0.140
9	IMPT	0.971	<0.01	0.02	0.036
	IMRT	0.972	0.02	0.02	0.130
10	IMPT	0.967	<0.01	0.02	0.050
	IMRT	0.973	0.01	0.02	0.120
11	IMPT	0.971	<0.01	0.02	0.060
	IMRT	0.972	0.01	0.02	0.080
12	IMPT	0.971	<0.01	0.02	0.036
	IMRT	0.972	0.02	0.02	0.100
13	IMPT	0.971	<0.01	0.02	0.042
	IMRT	0.971	0.01	0.02	0.110
14	IMPT	0.969	<0.01	0.02	0.069
	IMRT	0.972	0.03	0.02	0.230
15	IMPT	0.971	<0.01	0.02	0.068
	IMRT	0.971	0.01	0.02	0.150
16	IMPT	0.972	<0.01	0.02	0.073
	IMRT	0.97	0.02	0.02	0.170

TABLE 7.2: Transition probabilities calculated based on the dose to each patient. Duplicates are indicated by an asterisk.

7.6 Mean organ doses by patient

The mean radiation doses that would be received by each organ as a result of each treatment are summarised in Table 7.3. The difference in dose for each modality is listed

for each patient.

ORGAN DOSES

Patient ID	Lung			Heart		
	IMPT (Gy)	IMRT (Gy)	Δ (Gy)	IMPT (Gy)	IMRT (Gy)	Δ (Gy)
1	0.74	3.00	2.26	0.05	0.25	0.20
2	0.94	4.54	3.60	0.08	0.44	0.37
3	0.92	3.11	2.20	0.04	0.34	0.30
3*	1.29	2.92	1.63	0.06	0.34	0.28
4	0.89	7.95	7.06	0.03	0.64	0.62
5	0.91	5.24	4.34	0.03	0.58	0.55
6	1.41	3.83	2.42	0.20	0.34	0.14
7	1.16	6.83	5.67	0.06	1.23	1.17
8	1.32	4.66	3.34	0.06	0.47	0.41
8*	1.38	4.67	3.28	0.08	0.48	0.40
9	0.89	6.45	5.56	0.07	0.69	0.61
10	1.14	4.82	3.68	0.02	0.71	0.69
11	1.02	1.86	0.84	0.10	0.32	0.22
12	1.35	7.80	6.46	0.10	2.04	1.95
13	1.15	5.58	4.42	0.10	3.31	3.21
14	1.19	10.00	8.81	0.14	2.68	2.54
15	1.18	3.70	2.52	0.06	0.39	0.33
16	1.36	6.32	4.96	0.07	0.54	0.47

TABLE 7.3: Mean doses to the heart and lung for each patient and treatment modality. Duplicates are indicated by an asterisk.

7.7 Discussion and conclusion

The proton treatment cost-effectiveness was predicted for each member of the left-sided breast cancer patients in the cohort. However, the presented approach has limitations that should be considered when interpreting the results.

The estimation of costs of adverse events is inherently challenging and can be significantly more complex compared with the estimations employed here. For example, the cost of drugs is highly variable with time. In addition, productivity loss was estimated considering paid work only and work such as caring and volunteering was omitted. This could be particularly relevant for older patients who were assumed to have no potential for productivity loss and were predicted as having predominantly cost-ineffective proton treatments.

Another limitation is the uncertainty associated with the time for toxicities to develop. This may have resulted in uncertainties in the estimated cost and QALYs. Uncertainties in the dose were not considered as it was not possible to perform a robust analysis in the treatment planning system used to develop the plans used in this work. Furthermore, there was a lack of data available on the uncertainties associated with the radiobiological model parameters.

The publication presented in this chapter demonstrates the developed model as a potential tool for selecting patients for proton therapy. This is particularly important for cancer types that would not normally be considered as standard indications for the treatment.

Chapter 8

Conclusions and Future Work

8.1 Conclusions

The model developed in this work provides a prediction of the clinical outcome associated with a given radiotherapy treatment plan. Specifically, the difference in dose distribution evident in a comparison of treatment plans is translated to a difference in clinical outcome. As a result, it has the ability to estimate the relative benefit of a proton treatment compared to conventional X-ray radiotherapy on an individual patient basis. This prediction is of great value in the absence of data from clinical trials of proton therapy and allows the patients who are expected to receive the greatest benefit from this limited resource to be identified.

Several stages of model development have been detailed in this work. These included:

1. The initial model, which was evaluated with a Monte Carlo simulation. While this model was functional, it was limited by large computational times.
2. The model that was evaluated analytically, giving the exact solution with a significantly reduced computation time. This model also assumed that the probability of death due to cancer or injury was determined by the stage/grade rather than the time the cancer/injury had been present.

3. Incorporation of the effects of dose and radiobiological model parameter uncertainties on the model predictions. It was found that uncertainties in dose delivery made it difficult to conclude whether proton therapy would improve the clinical outcome of the base of skull chordoma (BOSCh) demonstration patient. However, when considering NTCP model parameter uncertainty only, a clear benefit was predicted for the demonstration patient if they were to be treated with protons. A framework for considering the effect of uncertainty in the quality of life utility weights was also developed. These weights are particularly important drivers of outcomes.
4. Incorporation of the effect of treatment cost-effectiveness to allow the model to select patients for proton therapy through balancing treatment cost and patient quality of life.
5. An investigation of the patients selected by the model to receive proton therapy, based on proton treatment cost-effectiveness. Two distinct cohorts were considered, one consisting of BOSCh patients and the other consisting of left-sided breast cancer patients (treated with the deep inspiration breath hold technique (DIBH)). The former is considered to be a standard indication for proton therapy, while the latter is not. It was found that all BOSCh patients could be treated with proton therapy cost-effectively. In contrast, the majority of the breast cancer patient cohort could not be treated cost-effectively. This was largely due to the greater risk of toxicity associated with the radiation treatment of BOSCh, while the toxicity rate associated with DIBH (particularly cardiovascular disease) was found to be low for both proton and photon treatments. However, it was also evident that the breast cancer patients who could be treated cost-effectively had an elevated lung dose associated with the planned photon treatment. Therefore, individualised approaches to patient selection may prove useful for this indication. Conversely, the results are supportive of proton therapy being adopted as a standard treatment for BOSCh.

The developed toolkit has the potential to serve as the basis of a patient selection system for proton therapy in clinical environments. This is particularly relevant in Australia as proton therapy facilities commence construction. Patient selection systems are valuable

in ensuring the efficient and equitable delivery of healthcare, both of which are high priorities in any society.

8.2 Future work

The prediction offered by the presented model provides a valuable approach to patient selection for proton therapy. However, there remains significant potential for further development of the model both before and after clinical implementation. Recommendations for future work are as follows:

- External model validation. This is an essential requirement before this type of model can be implemented clinically, as it allows the clinical utility to be evaluated and the accuracy of the predictions to be tested.
- The inclusion of additional predictors. Currently, the model input variables are the dose distributions for each organ and the tumour, along with the patient gender and age at the time of treatment. Demographic specific toxicity models, which could give information on baseline risks or histories of certain complications, are not included. In addition, further information regarding the tumour stage at the time of treatment and the presence or absence of concurrent treatment could lead to more informed predictions.
- Individualised medicine and precision medicine have been gaining increased attention in recent years [103]. These approaches involve individual-specific treatments where individual variability is taken into account. While the model developed in this work aims to achieve this goal, there remains the potential for future development as advances are made in the field of genomics and as biomarkers are identified.
- Several of the parameters used in the model were not clinically founded due to a lack of appropriate data. However, as an increasing number of patients receive proton therapy treatment, and their treatment outcomes are recorded in multi-institutional data registries, it will be possible to gather observations of clinical outcomes and to

adjust the prediction through a mechanism of gradual feedback. This is also true for the quality of life utility weights. Therefore, it is possible to continuously refine the parameters in the toxicity models [28].

- A major limitation of the proposed style of patient selection system is that two treatment plans are required as input for the model to produce a comparison. This corresponds to a greater time investment from clinic staff. Automated planning has been proposed to address this issue [104] by facilitating the automatic generation of robust proton plans with a dose mimicking algorithm. Combining similar work with the model presented here would assist in addressing this practical aspect of clinical implementation.
- As proposed by Langendijk *et al.* [28], a possible future application of the model that has been developed is to investigate the efficacy of proton therapy in the treatment of certain indications. For example, while the model predicted that the majority of patients in the cohort considered in the publication P3 would receive a clear benefit from proton therapy, not all patients were treated with proton therapy for various reasons. Observations of the clinical outcomes of these patients would make for an interesting comparison with the outcomes of those patients who did receive proton therapy. As the two groups were predicted to have similar outcomes, any differences in observations could potentially be attributed to the proton treatment. Similar studies could be conducted in the future to investigate the efficacy of proton therapy.

Bibliography

- [1] AIHW A. Cancer in Australia: an overview 2017. Cancer series. 2017;(101).
- [2] Podgorsak EB. Radiation oncology physics: A handbook for teachers and students. IAEA; 2005. p. 161.
- [3] Barton MB, Jacob S, Shafiq J, et al. Estimating the demand for radiotherapy from the evidence: A review of changes from 2003 to 2012. Radiotherapy and oncology. 2014;112(1):140–144.
- [4] Tobias C, Lawrence J, Born J, et al. Pituitary irradiation with high-energy proton beams a preliminary report. Cancer research. 1958;18(2):121–134.
- [5] Khan F. The physics of radiation therapy. Lippincott Williams & Wilkins; 2014. p. 524–526.
- [6] Khan F. The physics of radiation therapy. Lippincott Williams & Wilkins; 2014. p. 529–531.
- [7] Hall EJ. Intensity-modulated radiation therapy, protons, and the risk of second cancers. International Journal of Radiation Oncology* Biology* Physics. 2006;65(1):1–7.
- [8] Brady LW, Heilmann H, Molls M. New technologies in radiation oncology. Springer; 2006. p. 349–356.
- [9] Kardar L, Li Y, Li X, et al. Evaluation and mitigation of the interplay effects of intensity modulated proton therapy for lung cancer in a clinical setting. Practical Radiation Oncology. 2014;4(6):e259–e268.

- [10] Paganetti H, Niemierko A, Ancukiewicz M, et al. Relative biological effectiveness (RBE) values for proton beam therapy. *International Journal of Radiation Oncology* Biology* Physics*. 2002;53(2):407–421.
- [11] Britten RA, Nazaryan V, Davis LK, et al. Variations in the RBE for cell killing along the depth-dose profile of a modulated proton therapy beam. *Radiation Research*. 2012;179(1):21–28.
- [12] Underwood TS, McMahon SJ. Proton relative biological effectiveness (RBE): a multiscale problem. *The British Journal of Radiology*. 2018;92(1093):20180004.
- [13] Brady LW, Heilmann H, Molls M. *New technologies in radiation oncology*. Springer; 2006. p. 359–360.
- [14] Kemp R, Ramiscal R, Fodero L, et al. Assessment of proton beam therapy (PBT). MSAC application 1455.1, assessment report. Commonwealth of Australia, Canberra, ACT. 2017;.
- [15] Goitein M, Jermann M. The relative costs of proton and x-ray radiation therapy. *Clinical Oncology*. 2003;15(1):S37–S50.
- [16] Cox J. Design and implementation of clinical trials of ion beam therapy. In: Linz U, editor. *Ion beam therapy: fundamentals, technology, clinical applications*. Chapter 19. Springer Science & Business Media; 2011.
- [17] Verma V, Simone CB, Mishra MV. Quality of life and patient-reported outcomes following proton radiation therapy: a systematic review. *JNCI: Journal of the National Cancer Institute*. 2017;110(4):341–353.
- [18] Mishra MV, Aggarwal S, Bentzen SM, et al. Establishing evidence-based indications for proton therapy: an overview of current clinical trials. *International Journal of Radiation Oncology* Biology* Physics*. 2017;97(2):228–235.
- [19] Olsen DR, Bruland ØS, Frykholm G, et al. Proton therapy—a systematic review of clinical effectiveness. *Radiotherapy and Oncology*. 2007;83(2):123–132.
- [20] Brada M, Pijls-Johannesma M, De Ruyscher D. Proton therapy in clinical practice: current clinical evidence. *Journal of Clinical Oncology*. 2007;25(8):965–970.

-
- [21] Allen AM, Pawlicki T, Dong L, et al. An evidence based review of proton beam therapy: the report of ASTRO's emerging technology committee. *Radiotherapy and Oncology*. 2012;103(1):8–11.
- [22] Glimelius B, Montelius A. Proton beam therapy—do we need the randomised trials and can we do them? *Radiotherapy and Oncology*. 2007;83(2):105–109.
- [23] Suit H, Kooy H, Trofimov A, et al. Should positive phase iii clinical trial data be required before proton beam therapy is more widely adopted? No. *Radiotherapy and Oncology*. 2008;86(2):148–153.
- [24] Goitein M, Cox JD. Should randomized clinical trials be required for proton radiotherapy? *Journal of Clinical Oncology*. 2008;26(2):175–176.
- [25] Freedman B. Equipoise and the ethics of clinical research. *New England Journal of Medicine*. 1987;.
- [26] Dahl O. Protons. A step forward or perhaps only more expensive radiation therapy? *Acta Oncologica*. 2005;44(8):798–800.
- [27] Darby SC, Ewertz M, McGale P, et al. Risk of ischemic heart disease in women after radiotherapy for breast cancer. *New England Journal of Medicine*. 2013; 368(11):987–998.
- [28] Langendijk JA, Lambin P, De Ruyscher D, et al. Selection of patients for radiotherapy with protons aiming at reduction of side effects: the model-based approach. *Radiotherapy and Oncology*. 2013;107(3):267–273.
- [29] Lambin P, van Stiphout RG, Starmans MH, et al. Predicting outcomes in radiation oncology—multifactorial decision support systems. *Nature Reviews Clinical Oncology*. 2013;10(1):27–40.
- [30] Brodin NP, Maraldo MV, Aznar MC, et al. Interactive decision-support tool for risk-based radiation therapy plan comparison for Hodgkin lymphoma. *International Journal of Radiation Oncology* Biology* Physics*. 2014;88(2):433–445.

- [31] Smith WP, Doctor J, Meyer J, et al. A decision aid for intensity-modulated radiation-therapy plan selection in prostate cancer based on a prognostic Bayesian network and a Markov model. *Artificial Intelligence in Medicine*. 2009;46(2):119–130.
- [32] Delaney G, Jacob S, Featherstone C, et al. The role of radiotherapy in cancer treatment: estimating optimal utilization from a review of evidence-based clinical guidelines. *Cancer: Interdisciplinary International Journal of the American Cancer Society*. 2005;104(6):1129–1137.
- [33] Mee T, Kirkby N, Kirkby KJ. Mathematical modelling for patient selection in proton therapy. *Clinical Oncology*. 2018;30(5):299–306.
- [34] Grutters JP, Abrams KR, De Ruyscher D, et al. When wait for more evidence? Real options analysis in proton therapy. *The Oncologist*. 2011;.
- [35] Widder J, Van Der Schaaf A, Lambin P, et al. The quest for evidence for proton therapy: model-based approach and precision medicine. *International Journal of Radiation Oncology* Biology* Physics*. 2016;95(1):30–36.
- [36] Bekelman JE, Asch DA, Tochner Z, et al. Principles and reality of proton therapy treatment allocation. *International Journal of Radiation Oncology* Biology* Physics*. 2014;89(3):499–508.
- [37] Jakobi A, Bandurska-Luque A, Stützer K, et al. Identification of patient benefit from proton therapy for advanced head and neck cancer patients based on individual and subgroup normal tissue complication probability analysis. *International Journal of Radiation Oncology* Biology* Physics*. 2015;92(5):1165–1174.
- [38] Verma V, Shah C, Rwigema JCM, et al. Cost-comparativeness of proton versus photon therapy. *Chinese clinical oncology*. 2016;5(4).
- [39] Verma V, Mishra MV, Mehta MP. A systematic review of the cost and cost-effectiveness studies of proton radiotherapy. *Cancer*. 2016;122(10):1483–1501.

- [40] Lundkvist J, Ekman M, Ericsson SR, et al. Economic evaluation of proton radiation therapy in the treatment of breast cancer. *Radiotherapy and Oncology*. 2005; 75(2):179–185.
- [41] Lundkvist J, Ekman M, Ericsson SR, et al. Cost-effectiveness of proton radiation in the treatment of childhood medulloblastoma. *Cancer*. 2005;103(4):793–801.
- [42] Lundkvist J, Ekman M, Ericsson SR, et al. Proton therapy of cancer: potential clinical advantages and cost-effectiveness. *Acta oncologica*. 2005;44(8):850–861.
- [43] Mailhot Vega RB, Kim J, Hollander A, et al. Cost effectiveness of proton versus photon radiation therapy with respect to the risk of growth hormone deficiency in children. *Cancer*. 2015;121(10):1694–1702.
- [44] Mailhot Vega RB, Kim J, Bussière M, et al. Cost effectiveness of proton therapy compared with photon therapy in the management of pediatric medulloblastoma. *Cancer*. 2013;119(24):4299–4307.
- [45] Mailhot Vega RB, Ishaq O, Raldow A, et al. Establishing cost-effective allocation of proton therapy for breast irradiation. *International Journal of Radiation Oncology* Biology* Physics*. 2016;95(1):11–18.
- [46] Ramaekers BL, Grutters JP, Pijls-Johannesma M, et al. Protons in head-and-neck cancer: bridging the gap of evidence. *International Journal of Radiation Oncology* Biology* Physics*. 2013;85(5):1282–1288.
- [47] Grimmett G, Stirzaker D. *Probability and random processes*. Oxford university press; 2001. p. 226–231.
- [48] Abler D, Kanellopoulos V, Davies J, et al. Data-driven markov models and their application in the evaluation of adverse events in radiotherapy. *Journal of radiation research*. 2013;54(suppl_1):i49–i55.
- [49] Punglia RS, Burstein HJ, Weeks JC. Radiation therapy for ductal carcinoma in situ: a decision analysis. *Cancer*. 2012;118(3):603–611.
- [50] The assessment of new radiation oncology technologies and treatments (ANRO-TAT) project final report. Tasman Radiation Oncology Group (TROG) ; 2012.

- [51] Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. *Medical Decision Making*. 1993;13(4):322–338.
- [52] Stahl JE. Modelling methods for pharmacoeconomics and health technology assessment. *Pharmacoeconomics*. 2008;26(2):131–148.
- [53] Caro JJ, Möller J, Getsios D. Discrete event simulation: the preferred technique for health economic evaluations? *Value in health*. 2010;13(8):1056–1060.
- [54] Standfield L, Comans T, Scuffham P. Markov modeling and discrete event simulation in health care: a systematic comparison. *International journal of technology assessment in health care*. 2014;30(2):165–172.
- [55] Weinstein MC, Torrance G, McGuire A. QALYs: the basics. *Value in health*. 2009;12:S5–S9.
- [56] Kharroubi SA, Edlin R, Meads D, et al. Use of bayesian Markov chain Monte Carlo methods to estimate EQ-5D utility scores from EORTC QLQ data in myeloma for use in cost-effectiveness analysis. *Medical Decision Making*. 2015;35(3):351–360.
- [57] Chadwick K, Leenhouts H. A molecular theory of cell survival. *Physics in Medicine & Biology*. 1973;18(1):78.
- [58] Allen Li X, Alber M, Deasy JO, et al. The use and QA of biologically related models for treatment planning: Short report of the TG-166 of the therapy physics committee of the AAPM. *Medical Physics*. 2012;39(3):1386–1409.
- [59] Lind BK, Mavroidis P, Hyödynmaa S, et al. Optimization of the dose level for a given treatment plan to maximize the complication-free tumor cure. *Acta Oncologica*. 1999;38(6):787–798.
- [60] Lyman JT. Complication probability as assessed from dose-volume histograms. *Radiation Research*. 1985;104(2s):S13–S19.
- [61] Kutcher GJ, Burman C. Calculation of complication probability factors for non-uniform normal tissue irradiation: The effective volume method. *International Journal of Radiation Oncology* Biology* Physics*. 1989;16(6):1623–1630.

- [62] Burman C, Kutcher G, Emami B, et al. Fitting of normal tissue tolerance data to an analytic function. *International Journal of Radiation Oncology* Biology* Physics*. 1991;21(1):123–135.
- [63] De Marzi L, Feuvret L, Boulé T, et al. Use of gEUD for predicting ear and pituitary gland damage following proton and photon radiation therapy. *The British Journal of Radiology*. 2015;88(1048):20140413.
- [64] Lee TF, Yeh SA, Chao PJ, et al. Normal tissue complication probability modeling for cochlea constraints to avoid causing tinnitus after head-and-neck intensity-modulated radiation therapy. *Radiation Oncology*. 2015;10(1):194.
- [65] Bentzen SM, Constine LS, Deasy JO, et al. Quantitative analyses of normal tissue effects in the clinic (QUANTEC): an introduction to the scientific issues. *International Journal of Radiation Oncology* Biology* Physics*. 2010;76(3):S3–S9.
- [66] Beetz I, Schilstra C, Burlage FR, et al. Development of NTCP models for head and neck cancer patients treated with three-dimensional conformal radiotherapy for xerostomia and sticky saliva: the role of dosimetric and clinical factors. *Radiotherapy and Oncology*. 2012;105(1):86–93.
- [67] Beetz I, Schilstra C, van der Schaaf A, et al. NTCP models for patient-rated xerostomia and sticky saliva after treatment with intensity modulated radiotherapy for head and neck cancer: the role of dosimetric and clinical factors. *Radiotherapy and Oncology*. 2012;105(1):101–106.
- [68] Christianen ME, Schilstra C, Beetz I, et al. Predictive modelling for swallowing dysfunction after primary (chemo) radiation: results of a prospective observational study. *Radiotherapy and Oncology*. 2012;105(1):107–114.
- [69] Wopken K, Bijl HP, van der Schaaf A, et al. Development of a multivariable normal tissue complication probability (NTCP) model for tube feeding dependence after curative radiotherapy/chemo-radiotherapy in head and neck cancer. *Radiotherapy and Oncology*. 2014;113(1):95–101.

- [70] Dutz A, Lühr A, Agolli L, et al. Development and validation of NTCP models for acute side-effects resulting from proton beam therapy of brain tumours. *Radiotherapy and Oncology*. 2018;.
- [71] Beetz I, Schilstra C, van Luijk P, et al. External validation of three dimensional conformal radiotherapy based NTCP models for patient-rated xerostomia and sticky saliva among patients treated with intensity modulated radiotherapy. *Radiotherapy and Oncology*. 2012;105(1):94–100.
- [72] Blanchard P, Wong AJ, Gunn GB, et al. Toward a model-based patient selection strategy for proton therapy: external validation of photon-derived normal tissue complication probability models in a head and neck proton therapy cohort. *Radiotherapy and Oncology*. 2016;121(3):381–386.
- [73] Stokkevåg CH, Schneider U, Muren LP, et al. Radiation-induced cancer risk predictions in proton and heavy ion radiotherapy. *Physica Medica*. 2017;.
- [74] Council NR, et al. Health risks from exposure to low levels of ionizing radiation: BEIR VII phase 2. Vol. 7. National Academies Press; 2006.
- [75] Kaatsch P. Epidemiology of childhood cancer. *Cancer Treatment Reviews*. 2010; 36(4):277–285.
- [76] Mondlane G, Gubanski M, Lind PA, et al. Comparative study of the calculated risk of radiation-induced cancer after photon-and proton-beam based radiosurgery of liver metastases. *Physica Medica*. 2017;42:263–270.
- [77] Mu X, Björk-Eriksson T, Nill S, et al. Does electron and proton therapy reduce the risk of radiation induced cancer after spinal irradiation for childhood medulloblastoma? A comparative treatment planning study. *Acta Oncologica*. 2005;44(6):554–562.
- [78] Hall EJ, Wu CS. Radiation-induced second cancers: the impact of 3d-CRT and IMRT. *International Journal of Radiation Oncology* Biology* Physics*. 2003; 56(1):83–88.

- [79] Kry SF, Salehpour M, Followill DS, et al. The calculated risk of fatal secondary malignancies from intensity-modulated radiation therapy. *International Journal of Radiation Oncology* Biology* Physics*. 2005;62(4):1195–1203.
- [80] Chung CS, Yock TI, Nelson K, et al. Incidence of second malignancies among patients treated with proton versus photon radiation. *International Journal of Radiation Oncology* Biology* Physics*. 2013;87(1):46–52.
- [81] Yock TI, Caruso PA. Risk of second cancers after photon and proton radiotherapy: a review of the data. *Health Physics*. 2012;103(5):577–585.
- [82] Homann K, Howell R, Eley J, et al. The need for individualized studies to compare radiogenic second cancer (RSC) risk in proton versus photon Hodgkin Lymphoma patient treatments. *Journal of Proton Therapy*. 2016;1(1).
- [83] Timlin C, Warren D, Rowland B, et al. 3d calculation of radiation-induced second cancer risk including dose and tissue response heterogeneities. *Medical Physics*. 2015;42(2):866–876.
- [84] Schneider U, Sumila M, Robotka J. Site-specific dose-response relationships for cancer induction from the combined Japanese A-bomb and Hodgkin cohorts for doses relevant to radiotherapy. *Theoretical Biology and Medical Modelling*. 2011; 8(1):1–21.
- [85] Edge SBea. *AJCC cancer staging handbook: from the AJCC cancer staging manual*. Springer New York; 2010. p. 12.
- [86] Mayo C, Yorke E, Merchant TE. Radiation associated brainstem injury. *International Journal of Radiation Oncology* Biology* Physics*. 2010;76(3):S36–S41.
- [87] Debus J, Hug E, Liebsch N, et al. Brainstem tolerance to conformal radiotherapy of skull base tumors. *International Journal of Radiation Oncology* Biology* Physics*. 1997;39(5):967–975.
- [88] Rubin P, Constine LS, Marks LB. *Alert: Adverse late effects of cancer treatment, volume 2*. Springer-Verlag Berlin Heidelberg; 2014. Radiation Oncology.

- [89] Moore RE, Kearfott RB, Cloud MJ. Introduction to interval analysis. Vol. 110. Siam; 2009. p. 7–13.
- [90] Tengs TO, Wallace A. One thousand health-related quality-of-life estimates. *Medical Care*. 2000;38(6):583–637.
- [91] Qi XS, White J, Li XA. Is α/β for breast cancer really low? *Radiotherapy and Oncology*. 2011;100(2):282–288.
- [92] Okunieff P, Hoeckel M, Dunphy EP, et al. Oxygen tension distributions are sufficient to explain the local response of human breast tumors treated with radiation alone. *International Journal of Radiation Oncology• Biology• Physics*. 1993;26(4):631–636.
- [93] Seppenwoolde Y, Lebesque JV, De Jaeger K, et al. Comparing different NTCP models that predict the incidence of radiation pneumonitis. *International Journal of Radiation Oncology* Biology* Physics*. 2003;55(3):724–735.
- [94] Marks LB, Bentzen SM, Deasy JO, et al. Radiation dose–volume effects in the lung. *International Journal of Radiation Oncology* Biology* Physics*. 2010;76(3):S70–S76.
- [95] Sahle BW, Owen AJ, Mutowo MP, et al. Prevalence of heart failure in Australia: a systematic review. *BMC cardiovascular disorders*. 2016;16(1):32.
- [96] Santos AM, Marcu LG, Wong CM, et al. Risk estimation of second primary cancers after breast radiotherapy. *Acta Oncologica*. 2016;55(11):1331–1337.
- [97] Cardiovascular disease: Australian facts 2011. Cardiovascular disease series. Cat. no. CVD53. Australian Institute of Health and Welfare, Canberra: AIHW ; 2011.
- [98] Zheng H, Ehrlich F, Amin J. Productivity loss resulting from coronary heart disease in Australia. *Applied Health Economics and Health Policy*. 2010;8(3):179–189.
- [99] Australian Institute of Health and Welfare. Heart, stroke and vascular diseases- Australian facts 2004. AIHW Cat. no. CVD27. Canberra: AIHW and National Heart Foundation of Australia ; 2004. Available from: <http://www.aihw.gov.au/acim-books/>, Accessed: 3-11-2018.

-
- [100] Ghafoori P, Marks LB, Vujaskovic Z, et al. Radiation-induced lung injury. assessment, management, and prevention. *Oncology (Williston Park)*. 2008;22:37–47.
- [101] Item number 58503, Medicare Benefits Scheme ; 2018. Available from: <http://www9.health.gov.au/mbs/search.cfm?q=58503&sopt=S>, Accessed:5-11-2018.
- [102] Prednisone, Pharmaceutical Benefits Scheme ; 2018. Available from: <https://www.pbs.gov.au/medicine/item/1934T-1935W-1936X>, Accessed:5-11-2018.
- [103] Collins FS, Varmus H. A new initiative on precision medicine. *New England Journal of Medicine*. 2015;372(9):793–795.
- [104] Kierkels R, Fredriksson A, Both S, et al. Automated robust proton planning using dose-volume histogram-based mimicking of the photon reference dose and reducing organ at risk dose optimization. *International Journal of Radiation Oncology* Biology* Physics*. 2018;.