Supplementary Appendix

Adjuvant versus salvage radiotherapy following radical prostatectomy in high-risk prostate cancer: a decision analysis

Table of Contents

Supplementary Methods

Decision analytic tree

The decision analytic tree is presented in Supplementary Figure 1. Within each branch of the tree, tracker variables were used to assess complications including erectile dysfunction, urinary incontinence and bowel dysfunction.

Derivation of model probabilities and utilities

The model probabilities were derived from a review of the literature as detailed in the main body of the manuscript. A literature search of Medline was conducted as of November 1, 2015 searching from database inception. Combinations of the follow medical subject headings and text words and phrases were used: "radiotherapy", "adjuvant", "salvage", "survival", "recurrence", "metastasis", "randomized controlled trial", "follow-up studies", "complications", "incontinence", "erectile dysfunction", "impotence", and "bowel dysfunction". This search was supplemented by a hand literature search of the references of retrieved articles as well as an expert consensus panel comprising two radiation oncologists (G.M. and E.S.), three urologists (C.J.D.W., R.S., and R.K.N.), and an external methodologist (A.J., anesthesiologist). Further, we searched for prostate cancer specific utilities using combinations of the follow words and phrases: "decision model", "decision analysis", "decision making", "utility", "quality-adjusted life years" and "prostate cancer". Published review articles and previous decision analyses on prostate cancer were examined to ensure completeness.

Based on our literature review and expert consensus, the per-cycle (ie. monthly) probabilities used to populate our model are presented in Table 2. Unfortunately, for many of these probabilities, level 1 (randomized controlled trial) evidence does not exist to inform the model. For the probability of recurrence following surgery for patients in the salvage arm, the probability of recurrence following radiotherapy, the probability of metastasis follow recurrence, and the probability of prostate-cancer specific mortality following metastasis, we did not use a single citation to derive these figures. Instead, after a thorough literature review as described above, a multi-disciplinary panel comprising radiation oncologists, urologists and an external methodologist reviewed the available evidence and decided upon a composite value. The literature used, and estimates therein, are presented in Supplementary Table 1.

There was little evidence to inform complication rates between adjuvant and salvage radiotherapy. We assumed that post-operative radiotherapy did not increase the probability of erectile dysfunction or incontinence[16]. Based on evidence from the SWOG S8794 trial demonstrating that there was no difference in erectile dysfunction between patients with adjuvant radiotherapy and an observation strategy[3], we considered the probability of potency recovery and age-related erectile dysfunction to be the same in the two arms. Similarly, we assumed that the probability of incontinence was the same in the two arms based on observational data[16, 31]. Data on age-related erectile dysfunction was available for patients aged 40-69 based on a literature review[32]. Using the observed increases in erectile dysfunction from the literature, we extrapolated the probability of erectile dysfunction for patients over the age of 70.

We estimated the probability of having radical prostatectomy-related erectile dysfunction and incontinence at Cycle 0 based on literature-derived values at 3-months post-operative[12, 33]. Similarly, the utilities used to populate the model are presented in Table 1. We assumed, for the purposes of comparison between adjuvant and salvage radiotherapy strategies, that the utility of being status post-radical prostatectomy without radiotherapy or complication was 1. This is clearly not correct given the available literature. However, the disutility of this state would be present for all patients in the model, therefore, any deviation from this assumption would be non-informative from the model perspective. Clinically meaningful ranges have not been ascertained for many of these health states. As such, we employed the ranges available in the published literature.

Modelling details

We modeled each cycle as one month in duration given the relatively protracted natural history of prostate cancer. While others have used longer durations (one year[14]), we felt that one month allowed for greater detail with respect to the burden on active radiotherapy treatment as well as complications.

The lowest probability used to populate our model was 0.0013 corresponding to a 1 in 769 event. We sought to ensure that on each run of the model, this lowest probability event would occur a minimum of one time. Therefore, for our primary analysis, we performed 10,000 repetitions.

In keeping with modelling best practices, we performed a half-cycle correction for the utility associated with each health state.

Model calibration

Model calibration involves an iterative process of adjusting key model parameters in order to tune the decision model so that its output matches observed data. A calibration exercise involves identifying a source of observed data, determining which aspects of the observed data to calibrate the model toward (the targets), modifying the decision model so that it is able to report values corresponding to the calibration targets, determining which decision model parameters influence the value of the targets, creating sets of the selected model parameters, running the model with each set of parameters, and for each parameter set calculating a goodness-of-fit (GOF) score that provides a summary measure of how close the modelled target values are to the observed values.

Identification of observed data and targets

The decision model was calibrated to the randomized clinical trial (RCT) of Bolla et al. [4] of postoperative radiotherapy following radical prostatectomy (RP) versus a wait-and-see strategy. The arms of the Bolla RCT correspond to the ‘Salvage’ and ‘Adjuvant’ strategies compared in the decision model. Hereafter, we will refer to the names of the strategies as per the decision model nomenclature. We calibrated the decision model against the Salvage’ and ‘Adjuvant’ Kaplan-Meyer (KM) curves in the Bolla RCT corresponding to biochemical-progression-free survival (i.e. where the event of interest was the first of biochemical progression or death).

Calibration targets from the observed data consisted of the KM cumulative survival probabilities at two year increments from two through ten years after randomization for both the Salvage' and 'Adjuvant' arms. The observed survival probabilities were estimated graphically from the published KM curves. Since decision models follow hypothetical patients for their entire lives, and therefore have an unlimited time horizon, whereas RCTs have a limited follow-up time with censored observations, we estimated the censoring distribution for both arms of the Bolla trial as follows. At each two –year increment from 2 to 14 years, we estimated the probability of censoring from the KM survival probabilities and the number of patients at risk (indicated in the KM figures in the Bolla publication) as shown in Supplementary Table 2.

Where 'S(t)' is the cumulative survival probability (i.e. the probability of being free of biochemical progression or death), 'NaR' is the number at risk (from the KM figure), 'CumEventFree' is the number of patients free of the event of interest at each time point equal to NaR(0) * S(t), 'Events' is the number of events that occurred in the preceding interval equal to NaR(t-1)*[1-(S(t)/S(t-1))], 'LossToCohort' is equal to NaR(t-1) – NaR(t), 'Censored' is the number of censored events in the preceding interval equal to LossToCohort – Events, and 'Prop' is the proportion of the total group that were censored within each interval equal to Censored/NaR(0).

For each strategy, we estimated the weighted mean censoring time and its variance as:

$$\mu = \frac{\sum_{i=1}^{7}(Prop_{2i} * 2i)}{\sum_{i=1}^{7} Prop_{2i}}$$

$$\sigma^2 = \frac{\sum_{i=1}^{7} Prop_{2i} * (2i - \mu)^2}{\sum_{1=1}^{7} Prop_{2i}}$$

For the Salvage and Adjuvant arms the mean (standard deviation) censoring time estimates were 9.57 (3.13) and 9.97 (3.33) years, respectively. We then re-parameterized the estimated mean and standard deviation of the censoring times into the α and β parameters of gamma distributions as $\alpha = \mu^2/\sigma^2$ and $\beta = \mu/\sigma^2$.

Comparison of the observed proportion of censoring in the two arms of the Bolla RCT, and the proportion estimated from the corresponding gamma distributions are shown in Supplementary Figure 2.

Modification of the model to report calibration targets

In order to reflect the possibility of censored observations, we modified the decision model's stopping rules. For each hypothetical patient, we selected two random censoring times, one each from the gamma distributions for the 'Salvage' and 'Adjuvant' arms. Within each strategy, simulation for a hypothetical patient ended if he died or the elapsed simulation time was either more than 40 years or greater than the sampled censoring time.

For each parameter set (described below) we simulated 1000 post RP patients for each strategy. After each hypothetical patient's simulation, we wrote the observation time and an outcome indicator value to TreeAge global matrices 1 and 2 for the Salvage and Adjuvant arms,

respectively. The indicator value was set to '1' if either biochemical recurrence or death occurred and '0' is the hypothetical patient was censored. After the 1000 patients were simulated, we used the built-in TreeAge Python utility in TreeAge to sort global matrices 1 and 2 by observation time, calculate the 2,4,6,8, and 10 year cumulative survival probabilities, and store the latter in a row of global matrices 3 and 4 for the Salvage and Adjuvant strategies, respectively. With each new set of calibration parameters, we had the decision model write the associated survival probabilities into a separate row of global matrices 3 and 4.

Creating sets of model parameters that influence the value of the calibration targets in the decision model.

We *a priori* selected the following six parameters as being most likely to affect the calibration targets: the monthly probability of dying from prostate cancer, pDeath_CaP; an adjustment factor for the monthly probability of dying from other causes, pDeath_OC ; the monthly probability of prostate cancer recurrence after RP in the Salvage arm from 0 to 30 months, pRecurr_Salvg_lt30; the monthly probability of prostate cancer recurrence after RP in the Salvage arm beyond 30 months, pRecurr_Salvg_gt30; the relative risk of prostate cancer recurrence in the Adjuvant arm from 0 to 30 months RRrecurr_lt30; and the relative risk of prostate cancer recurrence in the Adjuvant arm beyond 30 months, RRrecurr_gt30. The 30 month time point represented a point of inflection in the Salvage KM curve in the Bolla RCT.

For each parameter, we selected two values below the base literature estimate, the base estimate itself, and two values above the base estimate in order to encompass a range from half to double the base estimate. The five values for each of the six parameters created 5^6 = 15,625 calibration parameter combinations. We ran the decision model with each of these sets which produced 15,625 rows of survival probabilities for the Salvage strategy in TreeAge global matrix 3 and the same number of rows for the Adjuvant strategy in global matrix 4. These matrices were exported and combined into a spreadsheet together with the associated parameter sets.

Calculation of the goodness-of-fit (GOF) score

For each parameter set, a Euclidian GOF was calculated across the KM survival probabilities for 2 through 10 years as:

$$GOF = \sqrt{\sum_{j=1}^{2}\sum_{i=1}^{5}\left(T_{2i,j} - M_{2i,j}\right)^2}$$

Where 'j' is the strategy number ('1' for Salvage, '2' for Adjuvant), $T_{2i,j}$ is the observed KM survival probability for the $2*i^{th}$ year and j^{th} strategy, and $M_{2i,j}$ is the corresponding modelled KM survival probability. Supplementary Figure 3 shows the Euclidian distance GOF result for the 15,625 parameter sets.

Averaging across the top 1% of parameter sets produced the calibrated parameter values (Supplementary Table 3).

Note that the monthly probability of death from other causes was obtained from mortality statistics from Statistics Canada and the pDeath_OC represents a multiplicative factor. Using the calibrated parameter values – the observed and modelled KM curves are shown in the manuscript as Figure 2.

These calibrated parameters were then used in subsequent two dimensional decision model simulations.

Two dimensional simulation

Two dimensional simulation (2D-sim) is a mode of running the decision model that captures both patient-level (first order) variability and parameter-level (second order) uncertainty. First order variability refers to the fact that individual hypothetical patients may vary with respect to characteristics such as the age at RP surgery and, given the current value of model parameters, the fact that individual patients may take a variable course through the model (i.e. have a variable life history). For example, given the monthly probability of incontinence, one patient may suffer the complication at a particular time point post RP while another patient may not. Second order uncertainty refers to the fact the model's parameters are themselves estimated with uncertainty from studies. For example, the monthly probability of incontinence may have a literature derived point estimate with surrounding confidence intervals.

In a 2D-sim analysis, parameter samples are drawn from distributions, and then with each sample, a number of hypothetical patients are simulated. In TreeAge nomenclature, each 2nd order parameter sample (each 'outer loop' iteration) is known as a 'sample' and each hypothetical patient (each 'inner-loop' iteration) is referred to as a 'trial'. During a course of a trial, the sampled hypothetical patient will run first through the 'Salvage' arm and then through the 'Adjuvant' arm. Within each strategy, the patient will experience a certain number of discounted, quality, adjusted, life months. The difference between the quality-adjusted life months experienced by the hypothetical patient in the two strategies represents the incremental benefit for that trial. The incremental benefits are averaged across the number of trials associated with each parameter sample. The average incremental benefits are then, in turn, averaged over the samples to provide the overall output from the model. In the current analysis, we also plotted the distribution of incremental benefit among the 2nd order sample iterations. For the current analyses, we simulated 10,000 outer loop samples and for each we simulated between 500 and 10,000 inner-loop trials. As the results were very stable, we present only the results of the analysis using 10,000 microsimulations.

The distributions we used for sampling and whether they were sampled on the outer loop (Per group of patients) or on the inner loop (Per patient) are shown in Supplementary Table 4.

Supplementary Results

External validation

We performed external validation by comparing oncologic outcomes derived from our model to estimates provided from the randomized controlled trials on this subject (Supplementary Table 5)[8-10]. We did not perform statistical tests for difference but assessed for comparability.

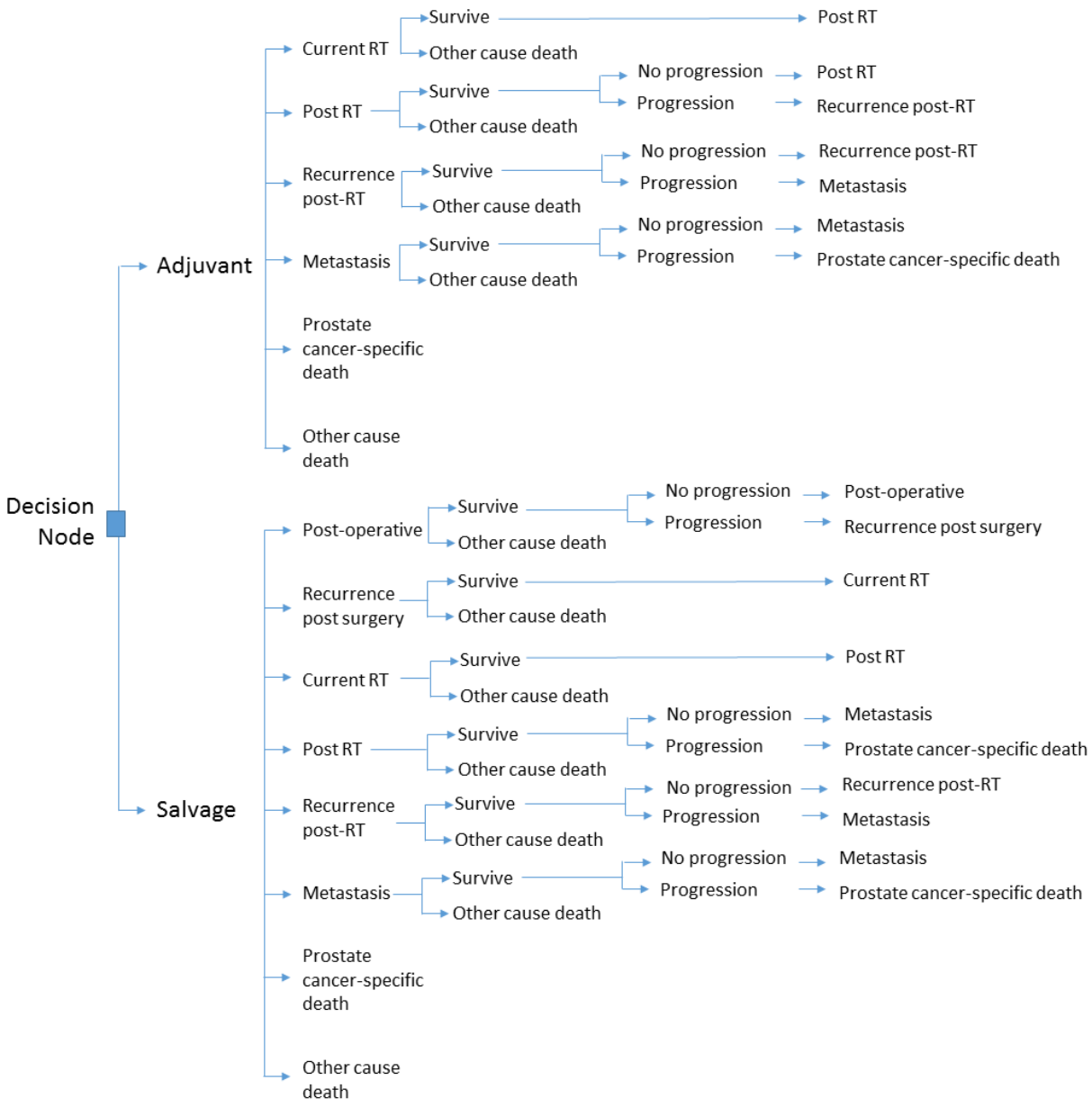
Overall life expectancy

When we examined overall life expectancy without adjustment for quality, there was a negligible difference in the two treatment strategies. The distribution of incremental benefit for the measures of overall life expectancy showed very little evidence of skew.

Similarly, when we performed this analysis without discounting, there was minimal difference between the two strategies. As with the analysis of overall life expectancy, there was little skew to this distribution.

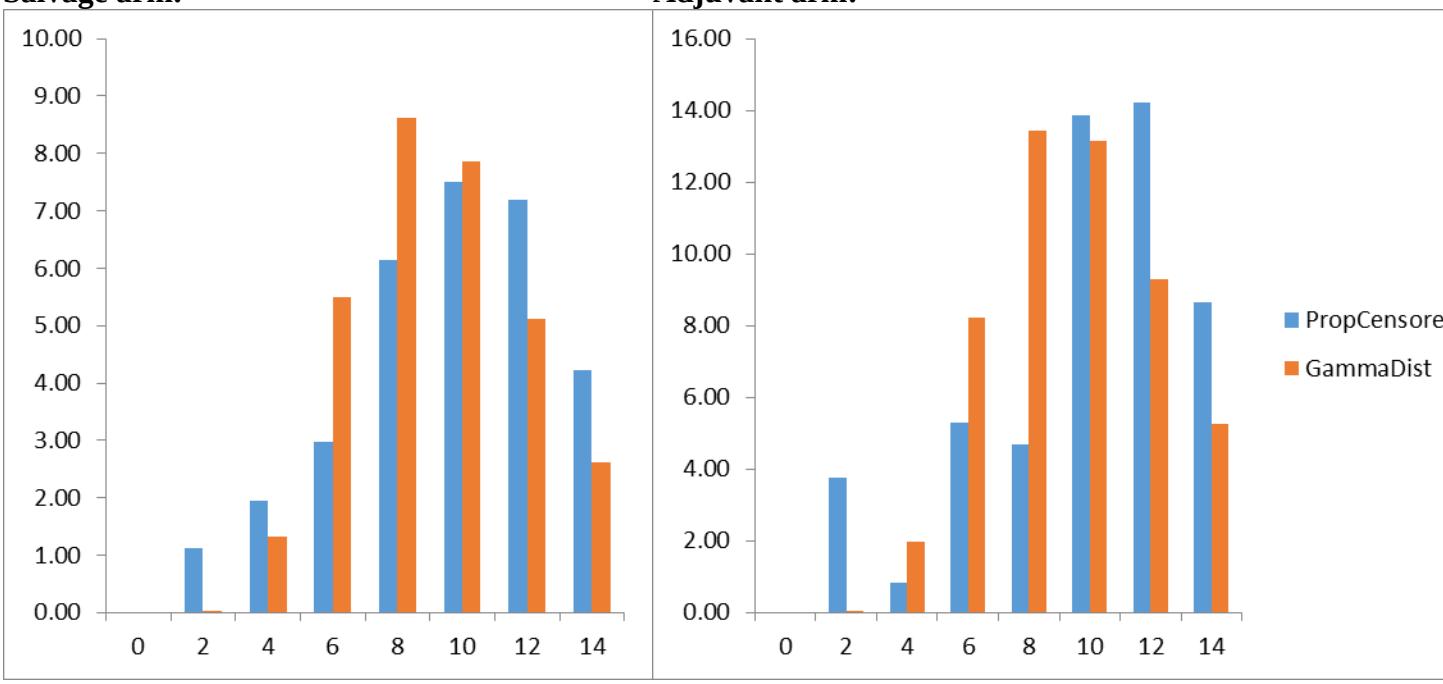
Supplementary Figures

Supplementary Figure 1. Decision analytic tree.

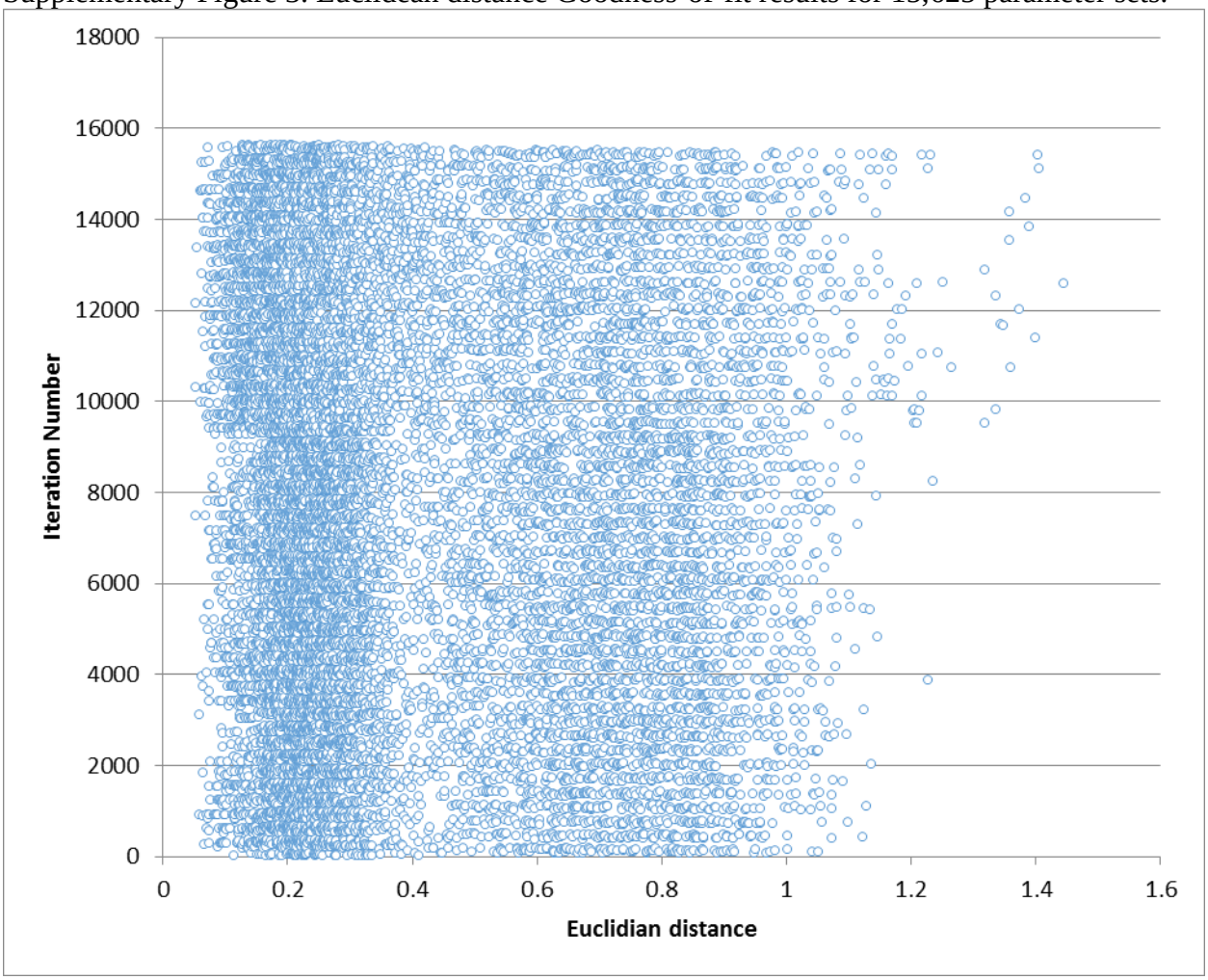


Supplementary Figure 2. Comparison of the proportion of censoring in the Bolla trial and the proportion estimated from the assigned gamma distributions.

Salvage arm: **Adjuvant arm:**



Supplementary Figure 3. Euclidean distance Goodness-of-fit results for 15,625 parameter sets.



Supplementary Tables

Supplementary Table 1. Literature used to derive estimates of oncologic outcomes to inform the Markov Monte Carlo model.

	Adjuvant intent			**Salvage intent**		
	Reference	**Probability**	**Per cycle prob**	**Reference**	**Probability**	**Per cycle prob**
RECURRENCE FOLLOWING RADICAL PROSTATECTOMY						
Option 1	NOT APPLICABLE			Bolla, Lancet 2012	42% at 4yr 50% at 6yr 55% at 10yr	0.011284 0.009581 0.006632
Option 2	NOT APPLICABLE			Thompson, JAMA 2006	55% at 5yr 70% at 10yr	0.01322 0.00998
Summary Result			N/A			0.010
RECURRENCE FOLLOWING RADIOTHERAPY						
Option 1	Trabulsi, Urology 2008	27% at 5yr	0.005231	Trabulsi, Urology 2008	50% at 5yr	0.011486
Option 2	Thompson, JAMA 2006	25% at 5yr 45% at 10yr	0.004783 0.004970	Stephenson, JAMA 2004	55% at 4yr 68% at 6yr	0.016498 0.015701
Option 3	Bolla, Lancet 2012	29% at 5yr 35% at 10yr	0.005692 0.003583			
Summary Result			0.005			0.015
METASTASIS FOLLOWING RECURRENCE AFTER RADIOTHERAPY						
Option 1	RTOG 9601	23% at 12yr	0.00181	RTOG 9601	23% at 12yr	0.00181
Option 2	Pound, JAMA 1999	35% at 5yr 55% at 10yr	0.00715 0.00663	Pound, JAMA 1999	35% at 5yr 55% at 10yr	0.00715 0.00663
Option 3	Roberts, Mayo Clin 2001	6% at 5yr 9% at 10yr	0.00103 0.00079	Roberts, Mayo Clin 2001	6% at 5yr 9% at 10yr	0.00103 0.00079
Summary Result			0.0018			0.0018
PROSTATE CANCER SPECIFIC MORTALITY FOLLOWING METASTASIS AFTER RADIOTHERAPY						
Option 1	American Cancer Society	28% at 5yr	0.00546	American Cancer Society	28% at 5yr	0.00546
Option 2	Cancer Research UK	30% at 5yr	0.00593	Cancer Research UK	30% at 5yr	0.00593
Option 3	Canadian Cancer Society	31% at 5yr	0.00617	Canadian Cancer Society	31% at 5yr	0.00617

Summary Result		0.00585			0.00585

Supplementary Table 2. Censoring probabilities derived from Bolla et al.[8].

Salvage							
Time	S(t)	NaR	CumEventFree	Events	LossToCohort	Censored	Prop
0	1	503	503	0	0	0	0
2	0.752688	373	379	124.3978	130	5.602151	0.011137
4	0.580645	278	292	85.25714	95	9.742857	0.019369
6	0.505376	227	254	36.03704	51	14.96296	0.029747
8	0.451613	172	227	24.14894	55	30.85106	0.061334
10	0.419355	122	211	12.28571	50	37.71429	0.074979
12	0.306452	53	154	32.84615	69	36.15385	0.071876
14	0.215054	16	108	15.80702	37	21.19298	0.042133
Sum				330.7799	487	156.2201	0.310577
Adjuvant							
Time	S(t)	NaR	CumEventFree	Events	LossToCohort	Censored	Prop
0	1	502	502	0	0	0	0
2	0.913978	440	459	43.1828	62	18.8172	0.037484
4	0.806452	384	405	51.76471	56	4.235294	0.008437
6	0.725806	319	364	38.4	65	26.6	0.052988
8	0.645161	260	324	35.44444	59	23.55556	0.046923
10	0.602151	173	302	17.33333	87	69.66667	0.138778
12	0.537634	83	270	18.53571	90	71.46429	0.142359
14	0.462366	28	232	11.62	55	43.38	0.086414
Sum				216.281	474	257.719	0.512364

Supplementary Table 3. Calibrated parameter estimates.

Parameter	**Base (literature) estimate**	**Calibrated value**
pDeath_CaP	0.00585	0.00802
pDeath_OC	1†	0.99038
pRecurr_Salvg_lt30	0.01	0.01317
pRecurr_Salvg_gt30	0.01	0.00479
Rrrecurr_lt30	0.5	0.42885
Rrrecurr_gt30	0.5	0.43910

Supplementary Table 4. Distributions used for sampling each variable in the decision model.

Name	Distribution type	Mean or LB	SD or UB	Sampling Rate
Disutility of bowel dysfunction	Beta	0.71	0.26	Per group of patients (2nd order)
Disutility of impotence	Beta	0.89	0.16	Per group of patients (2nd order)
Disutility of incontinence	Beta	0.83	0.13	Per group of patients (2nd order)
Probability of bowel dysfunction	Beta	0.001343	0.002797	Per patient (1st order)
Probability of post-op impotence	Beta	0.77	0.0384	Per patient (1st order)
Probability of regaining impotence	Uniform	0	0.1	Per patient (1st order)
Age-related probability of impotence				
Ages 40 - 49	Beta	0.010972	0.001908	Per patient (1st order)
Ages 50 - 59	Beta	0.029055	0.003868	Per patient (1st order)
Age >= 60	Beta	0.050641	0.008377	Per patient (1st order)
Probability of post RTx impotence	Uniform	0	1	Per patient (1st order)
Probability of post-op incontinence	Uniform	0	1	Per patient (1st order)
Probability of regaining continence	Uniform	0	0.05	Per patient (1st order)
Probability of CaP recurrence within 30 mos of RP	Beta	0.013167	0.000885	Per group of patients (2nd order)
Probability of CaP recurrence beyond 30 mos of RP	Beta	0.004785	0.000885	Per group of patients (2nd order)
Probability of CaP recurrence after salvage RTx	Beta	0.016498	0.001166	Per group of patients (2nd order)
Age at RP (years)	Gamma	65	5.9	Per patient (1st order)
Utility on RTx	Beta	0.73	0.29	Per group of patients (2nd order)
Utility with metastatic RP	Beta	0.25	0.11	Per group of patients (2nd order)
Utility after prior RTx	Beta	0.78	0.29	Per group of patients (2nd order)
Utility of local CaP recurrence	Beta	0.68	0.26	Per group of patients (2nd order)

Note: RTx = radiotherapy; CaP = prostate cancer; RP = radical prostatectomy

Supplementary Table 5. Five-year biochemical recurrence rates, based on literature review.

	Adjuvant radiotherapy	'Wait-and-see'
Bolla et al.	21.4 % (95% CI 16.4-26.3%)	44.2% (95% CI 38.3-50.0%)
Thompson et al.	~25%	~60%
Wiegel et al.	28% (98% CI 19-35%)	46% (95% CI 37-55%)