Estimating Transition Probabilities for a Three State Markov Model from Published Kaplan-Meier Curves

Eldon Spackman, PhD

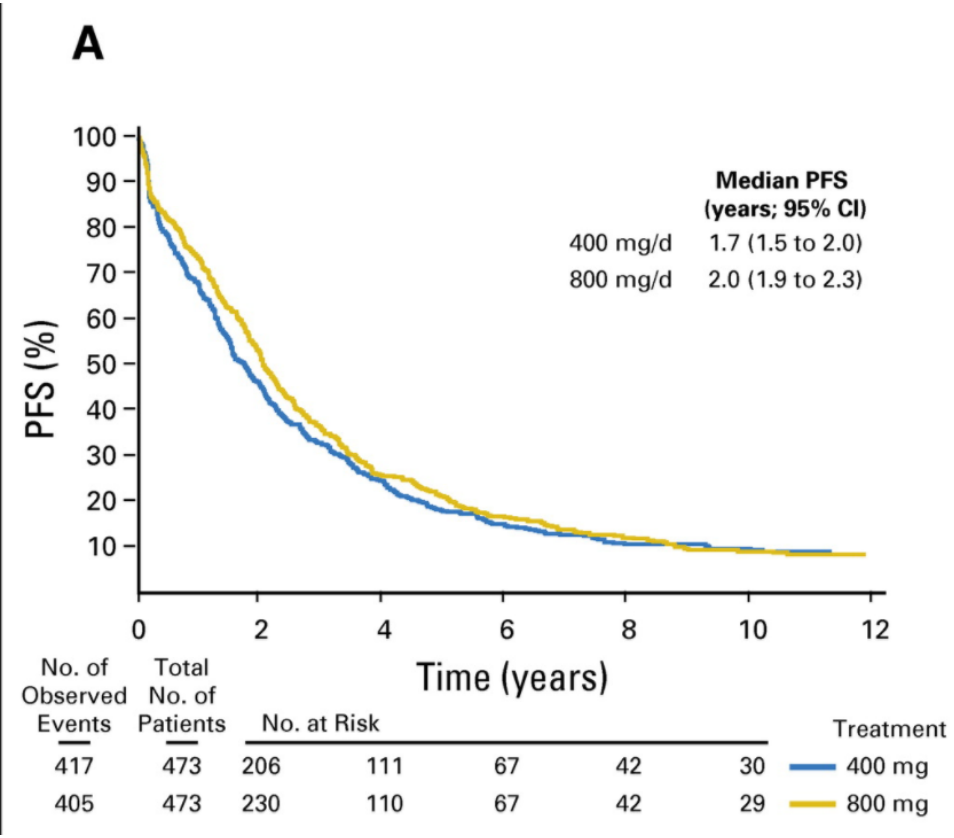
Assistant Professor

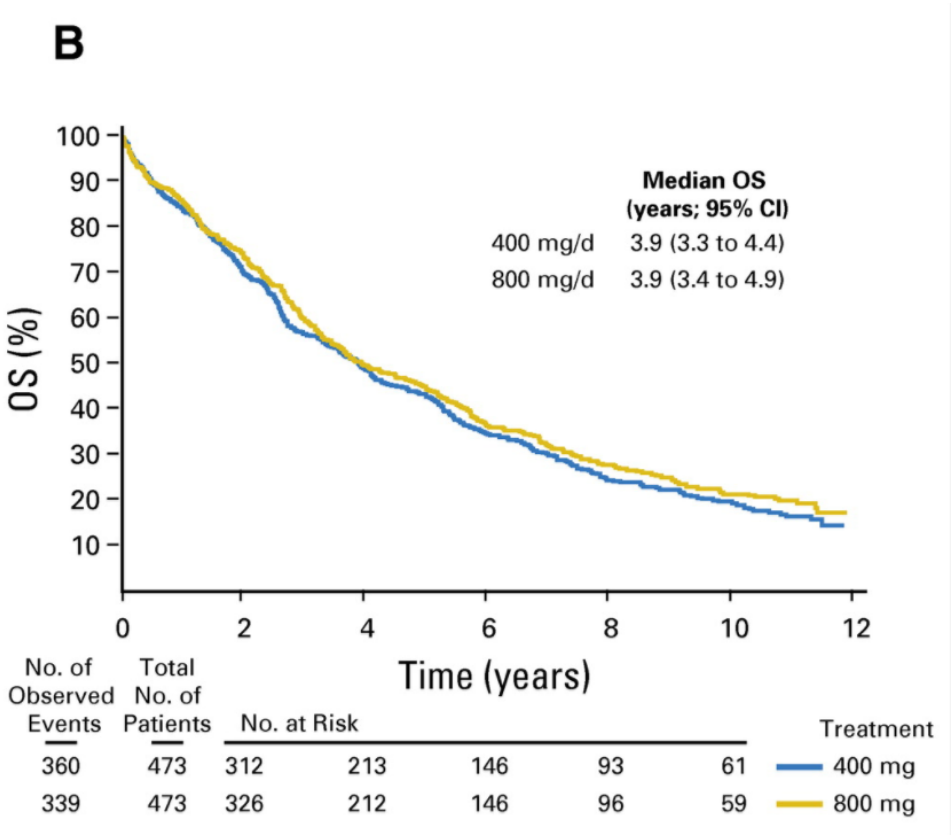
UNIVERSITY OF CALGARY
CUMMING SCHOOL OF MEDICINE
Department of Community Health Sciences

Contributors

- Colin Weaver, PhD student: University of Calgary
- Vishva Danthurebandara, PhD: University of Calgary
- Rita Faria, MSc: CHE, University of York
- Marta Soares: CHE, University of York

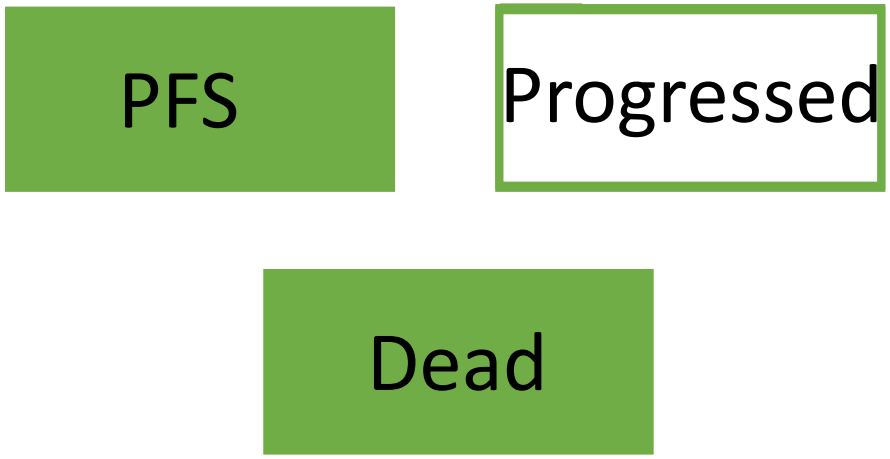
The Problem

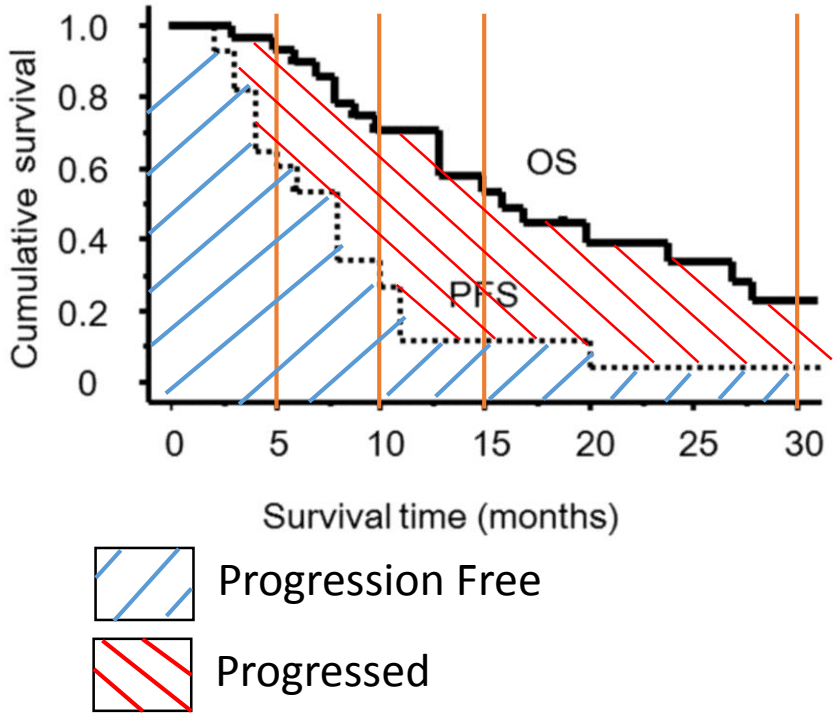




Casali PG, et al. J Clin Oncol. 2017:JCO2016710228. doi: 10.1200/JCO.2016.71.0228

Cost-Effectiveness Model: Partition Survival Analysis





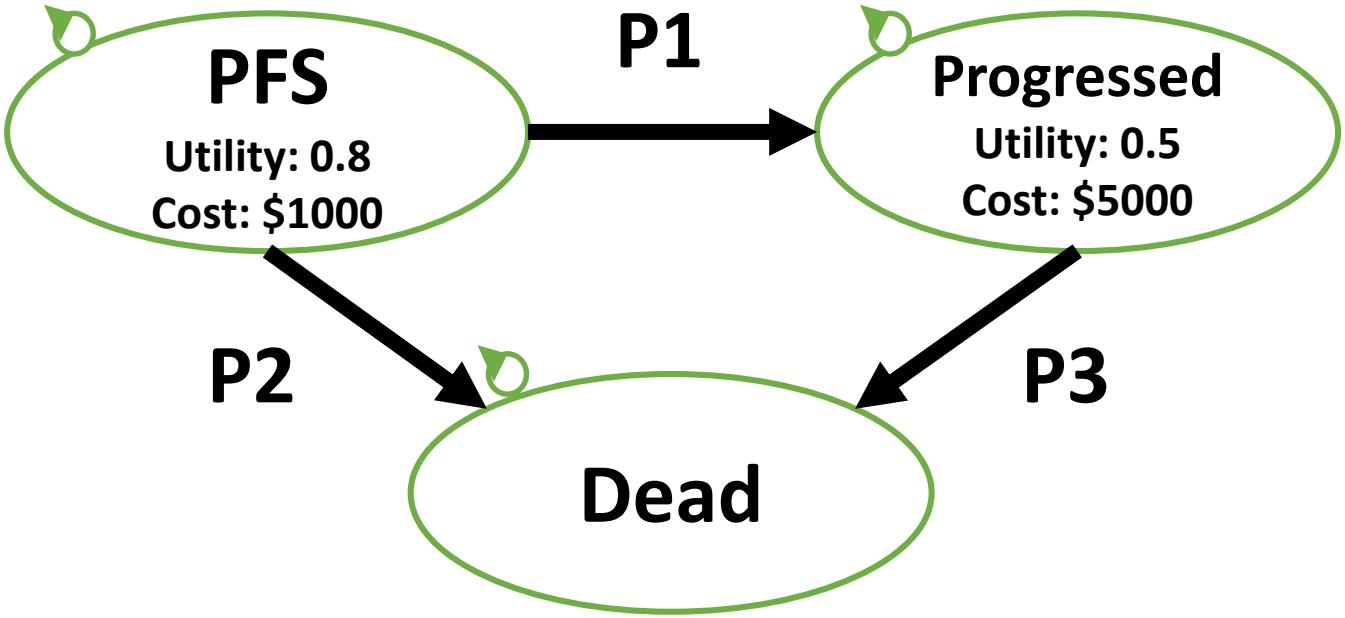
The most commonly used decision modelling approach for appraisals of interventions for advanced or metastatic cancers.

Woods B, et al. NICE DSU Technical Support Document 19. Partitioned Survival Analysis for Decision Modelling in Health Care: A Critical Review. 2017 [Available from http://www.nicedsu.org.uk]

The Challenge of Partition Survival Analysis

- Can not separate OS and PFS
  - Can not test differential treatment effects to different components of the disease process
- Limited sensitivity of extrapolated results
- Combining different data sources

Cost-Effectiveness Model: State Transition



10 cycles
1.5% discounting per cycle

The Challenge of State Transition Modelling

- The main challenge in the use of state transition models relates to the estimation of the required transition probabilities.
  - Particularly when only summary data are available

- Two Potential Methods for Estimating Transition Probabilities from Kaplan-Meier Curves
  - Simulation
  - Optimization

Woods B, et al. NICE DSU Technical Support Document 19. Partitioned Survival Analysis for Decision Modelling in Health Care: A Critical Review. 2017 [Available from http://www.nicedsu.org.uk]

Simulation Method

- Digitize the Kaplan-Meier Curve (WebPlotDigitizer)
- Estimate 95% confidence intervals
- Test combinations of P1, P2 and P3 that stay within the 95% CI
  - Limitations: P2<P3, from 0.01 to 0.50 in 0.01 increments (62,500 possible combinations)
- Four methods for combining simulations
  - Mean
  - Median
  - Probabilistic
  - Best fit i.e. lowest confidence interval

Optimization Method

- Non-linear minimization on sum of the squared differences
  - *optim* function in R and the Nelder and Mead method
  - 20 random initial transition matrices were used to avoid finding local minima
  - Chose the matrix with the global minimum
  - Rejected any matrix where P2>P3

Comparing Simulation to Optimization

- Specify the true transition matrix
  - P1= 0.20, P2=0.05 and P3=0.20
- Generate individual patient data for 100 patients
  - Draws from random uniform distribution to determine transitions
- Estimate Kaplan-Meier curves and confidence intervals from IPD, accounting for censoring and using log-log transformation
- Apply simulation and optimization methods to the KM curves and CI
- Estimate transition probabilities from IPD for comparison
  - Transitions of interest divided by the number of individuals in the original state
- Compared outcomes using the mean, mean absolute error and the mean squared error

Results

- Simulation
  - Calculated the number of matrices using a grid of 0.01 increments for 1000 K-M curves

Minimum	1st	Median	Mean	3rd	Maximum
0 (18 trials)	275	406	417.3	553	1234

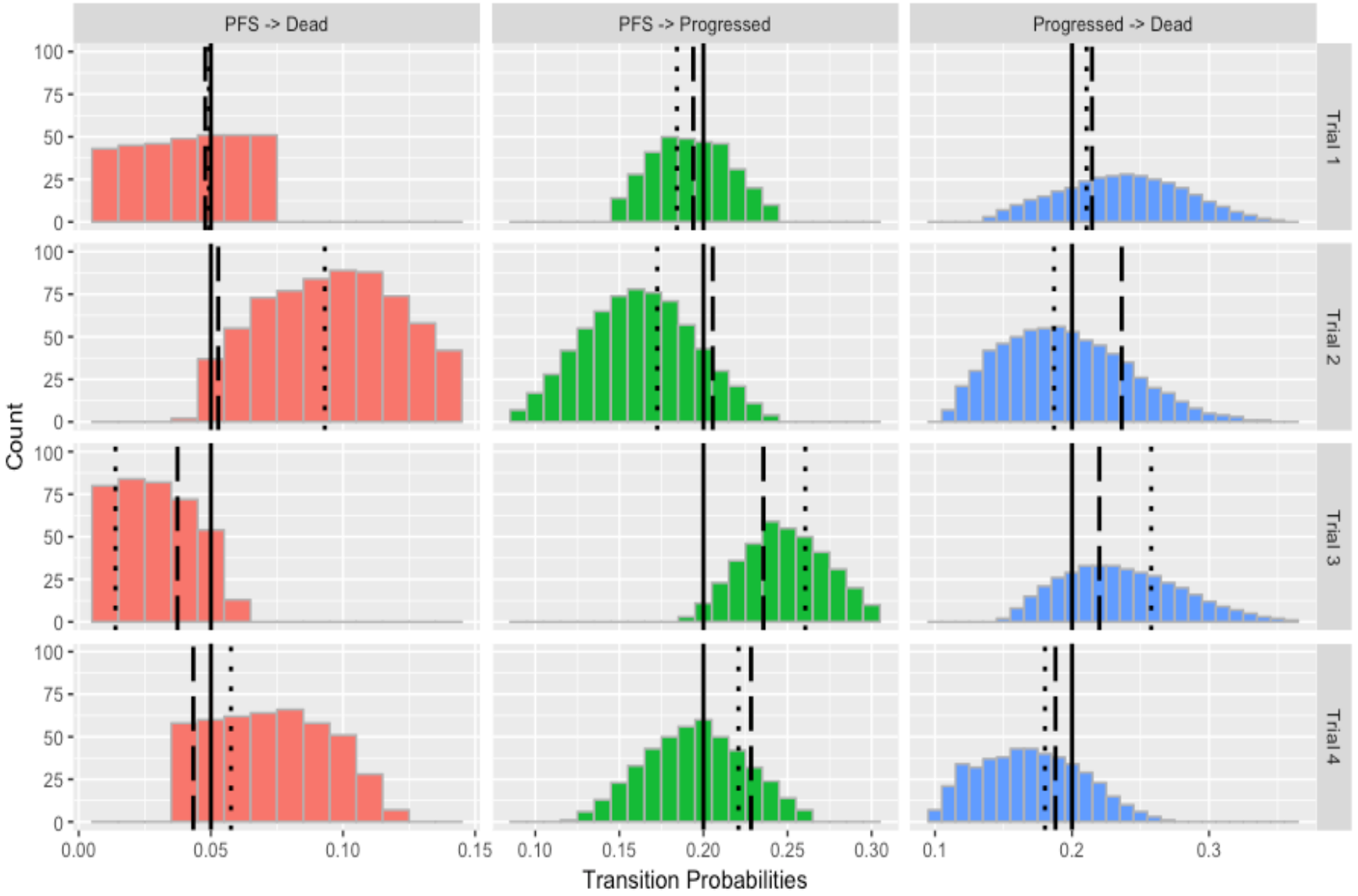
5.6% <100, excluded from future analysis

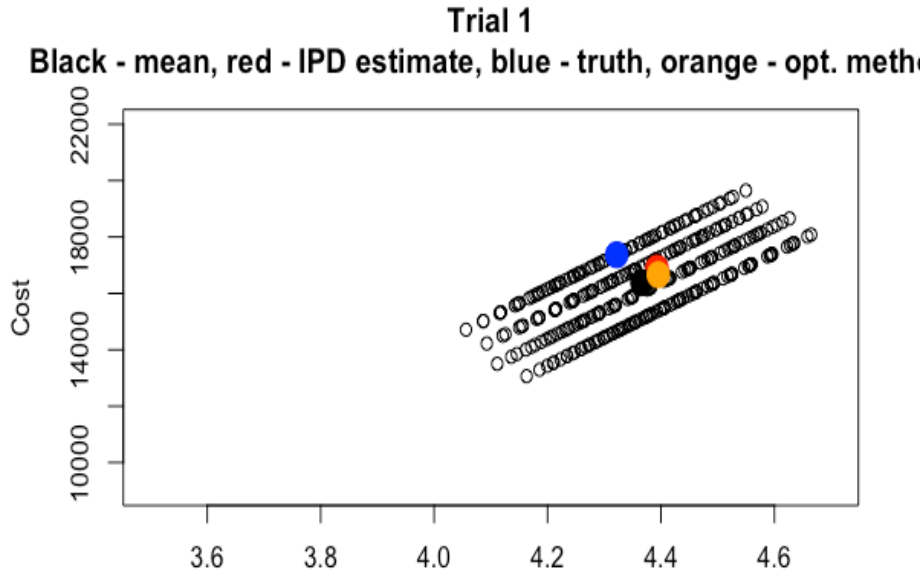
- Optimization
  - Tested the number of random starting values
    - Tested 1000 trials to see if 20 is sufficient
  - Most trials had 3-5 minima within 10%

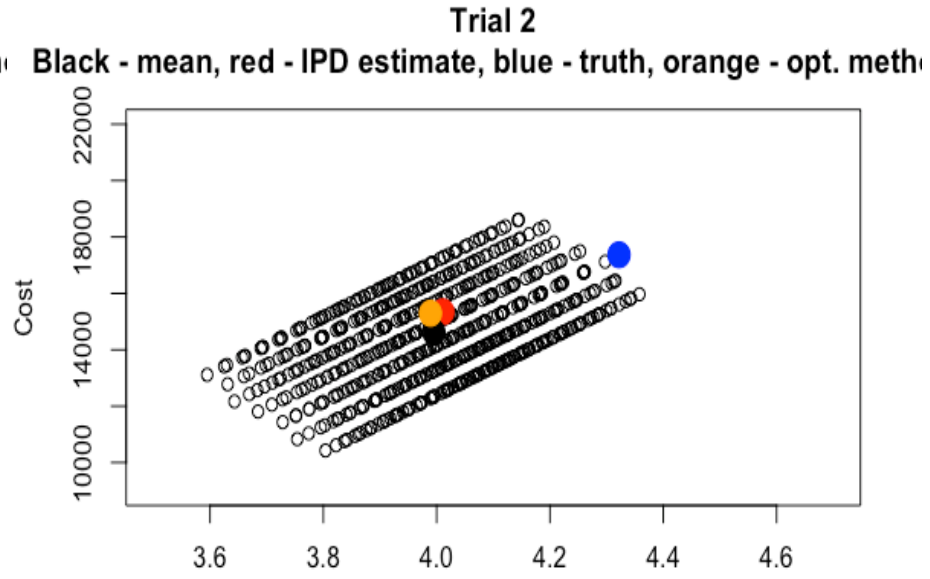
# Minima	1	2	3	4	5	6
# Trials	45	91	452	230	138	44

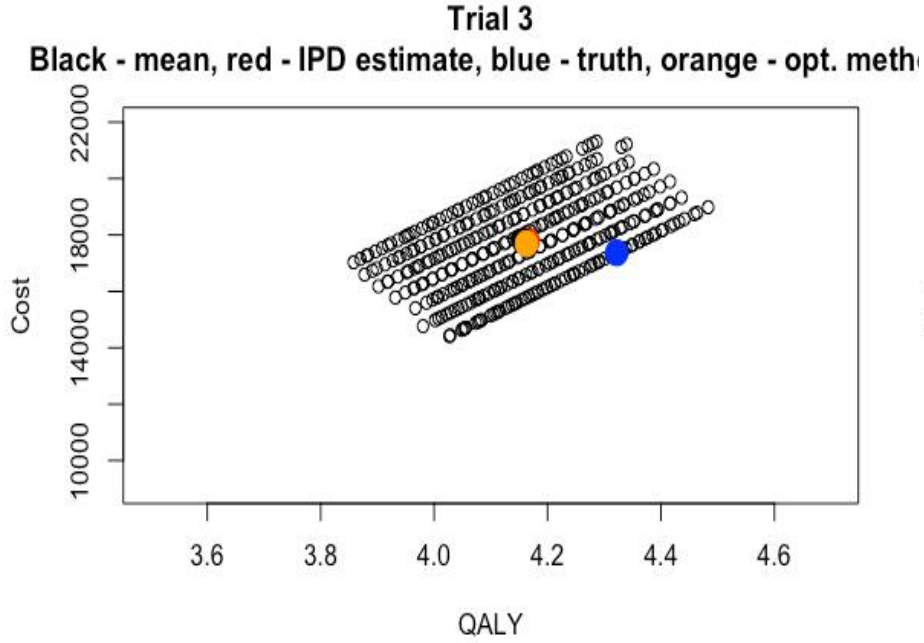
999 of 1000 trials had the same minima using 5 starting values compared to 20

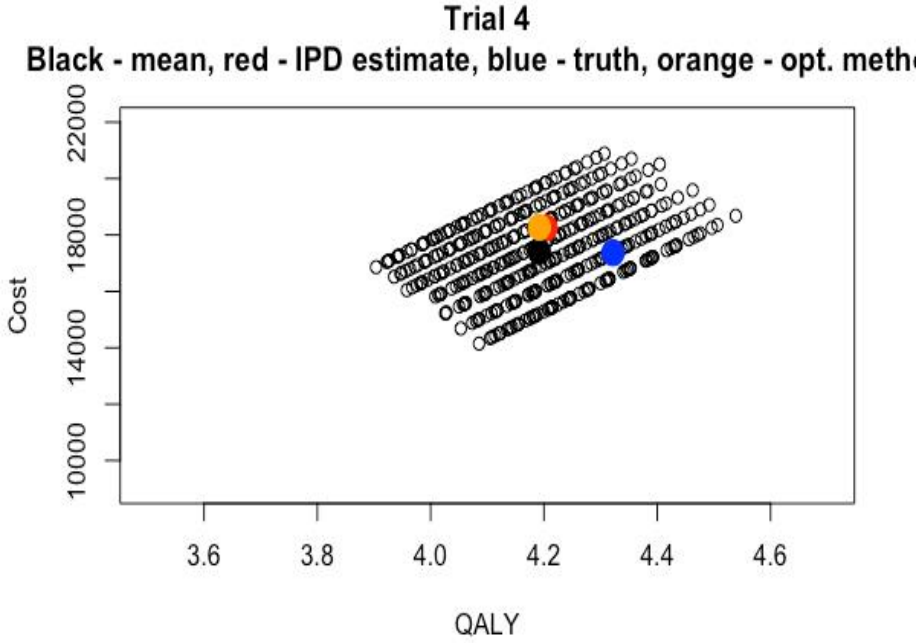
Results: Estimated transition probabilities (Trial 1: 336 , Trial 2: 679, Trial 3: 385, Trial 4: 454)



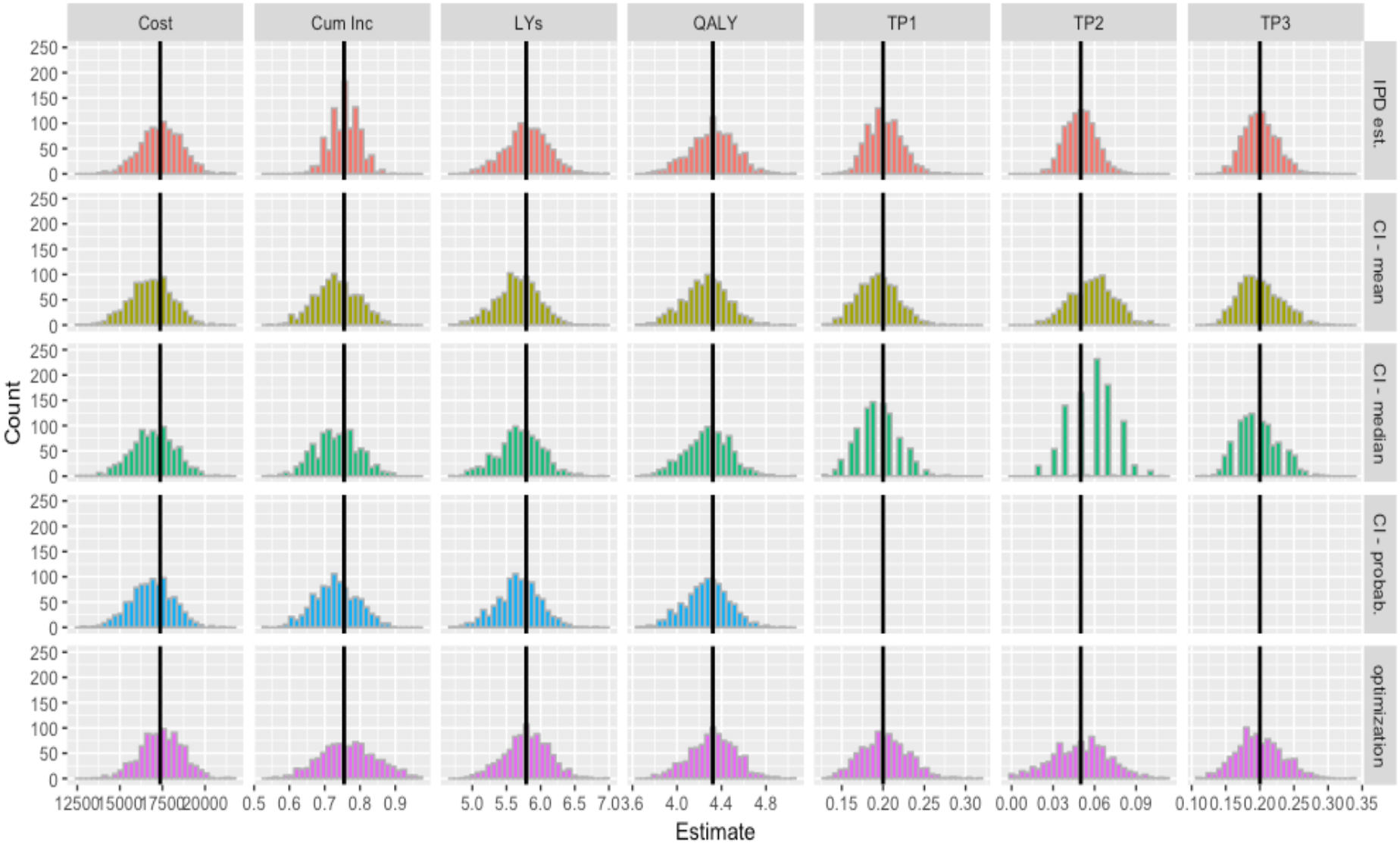

Trial 1
Black - mean, red - IPD estimate, blue - truth, orange - opt. meth
Cost
10000
14000
18000
22000
3.6
3.8
4.0
4.2
4.4
4.6


Trial 2
Black - mean, red - IPD estimate, blue - truth, orange - opt. meth
Cost
10000
14000
18000
22000
3.6
3.8
4.0
4.2
4.4
4.6


Trial 3
Black - mean, red - IPD estimate, blue - truth, orange - opt. meth
Cost
10000
14000
18000
22000
3.6
3.8
4.0
4.2
4.4
4.6
QALY


Trial 4
Black - mean, red - IPD estimate, blue - truth, orange - opt. meth
Cost
10000
14000
18000
22000
3.6
3.8
4.0
4.2
4.4
4.6
QALY

Results: Estimated transition probabilities and outcomes of 1000 trials



		Truth	Estimate from IPD	Within 95% CIs		Optimization
				Mean	Median	
TP 1, PFS->P	Mean	0.20	0.202	0.194	0.194	0.202
	MAE		0.017	0.020	0.020	0.023
	MSE		0.0004	0.0006	0.0006	0.0008
TP 2, PFS->D	Mean	0.05	0.050	0.059	0.058	0.050
	MAE		0.009	0.015	0.015	0.017
	MSE		0.0001	0.0003	0.0003	0.0004
TP 3, P->D	Mean	0.20	0.201	0.200	0.198	0.200
	MAE		0.019	0.025	0.025	0.027
	MSE		0.0006	0.0009	0.0009	0.0012

TP – transition probability, PFS – progression free survival, P - progressed, D – dead
IPD – individual patient data, MAE – mean absolute error, MSE – mean squared error

		Truth	Estimate from IPD	Within 95% CI			Optimization
				Mean	Median	Probabilistic	
Exp. LYs	Mean	5.79	5.79	5.66	5.70	5.68	5.79
	MAE		0.25	0.27	0.27	0.26	0.26
	MSE		0.10	0.11	0.11	0.11	0.10
Cum.	Mean	0.755	0.756	0.734	0.736	0.732	0.767
Inc.	MAE		0.033	0.052	0.053	0.053	0.065
	MSE		0.002	0.004	0.004	0.004	0.006
Costs	Mean	17 369	17424	16835	16969	16873	17429
	MAE		963	1081	1063	1061	983
	MSE		1442813	1849963	1770104	1776486	1499526
QALYs	Mean	4.32	4.32	4.26	4.28	4.27	4.32
	MAE		0.160	0.164	0.168	0.162	0.163
	MSE		0.041	0.043	0.044	0.042	0.042

Exp. LYs – expected life-years, Cum. Inc. – cumulative incidence, QALYs – quality adjusted life-years, IPD – individual patient data, MAE – mean absolute error, MSE – mean squared error

Advantages and Disadvantages of Simulation Method

- Advantages
  - Captures all options within CI
  - Matrices can be used directly in probabilistic analysis
- Disadvantages
  - May not always be able to find matrix within CI
    - Do transition probabilities have to produce curves within CI at all time points?
  - Must choose transition probabilities for deterministic analysis: mean, median, tightest CI

Advantages and Disadvantages of Optimization Method

- Advantages
  - Minimizing the distance between the actual and expected Kaplan-Meier curves is what we would intuitively like to minimize
- Disadvantages
  - No measure of uncertainty (no probabilistic analysis)
  - Does not weight the minimization of the curve by certainty, e.g. at later time points KM are more uncertain, might want to weight sum of squared difference at later time points less.

Next Steps

- Vary the simulation scenario (true transition probabilities, number of cycles, costs, QALYs)
- Consider transition probabilities that change over time
- Use additional information, cumulative incidence curve, censoring, hazard ratio
- Test use with recently published RCTs

Are there better methods?

- Are there other Bayesian methods to generate ‘pseudo IPD’?
- Would discrete event simulation be better?
  - Could we estimate to event more easilty/accurately?