Web-based Supplementary Materials for

Improving efficiency in clinical trials using auxiliary information; Application of a multi-state cure model

by A.S.C. Conlon[1], J.M.G. Taylor[1], D.J. Sargent[2]

[1]Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, 48109, U.S.A.,

[2]Mayo Clinic, Rochester, MN, USA

Web Appendix A: Details of 12 colon cancer trials.

Web Table 1: Data Summary of 12 colon cancer clinical trials

Study	N	Recurrences	Recurrence Without Death	Death Without Recur	Total Deaths	Longest Follow-Up (years)	% in Treatment Group
1	247	116	14	13	115	9.9	49.0%
2	408	139	11	44	172	9.1	62.5%
3	926	377	31	76	422	11.4	49.4%
4	914	380	36	106	450	9.9	75.2%
5	878	297	33	74	338	12.6	49.8%
6	724	275	10	132	397	13.2	48.2%
7	683	206	32	129	303	12.9	50.1%
8	1040	356	36	67	387	9.7	49.8%
9	2077	605	57	176	724	9.4	66.7%
10	2128	574	66	192	700	10.3	49.8%
11	1549	394	71	115	438	8	50.3%
12	2409	627	189	106	544	6	49.8%

Web Appendix B: Efficiency gains obtained through the proposed method utilizing the multi-state cure model: model based estimates.

Web Table 2: Kaplan-Meier treatment effect estimates (standard errors) and multi-state model estimates (posterior standard deviations) for 10 colon cancer trials

		$\Delta S(5)^*$		$\Delta DFS(3)^{**}$	
		Full Follow-up	Reduced Follow-up	Full Follow-up	Reduced Follow-up
Trial 1	Kaplan-Meier	0.090 (0.062)	0.098 (0.070)	0.122 (0.062)	0.132 (0.071)
	Full Multi-State Model	0.105 (0.049)	0.107 (0.057)	0.150 (0.051)	0.141 (0.057)
Trial 2	Kaplan-Meier	0.057 (0.049)	0.039 (0.049)	0.056 (0.047)	0.076 (0.050)
	Full Multi-State Model	0.051 (0.042)	0.053 (0.043)	0.068 (0.042)	0.070 (0.042)
Trial 4	Kaplan-Meier	-0.023 (0.037)	0.027 (0.043)	0.004 (0.037)	-0.003 (0.042)
	Full Multi-State Model	-0.024 (0.032)	-0.010 (0.027)	-0.015 (0.034)	-0.003 (0.032)
Trial 5	Kaplan-Meier	-0.023 (0.031)	-0.009 (0.035)	-0.015 (0.031)	-0.026 (0.035)
	Full Multi-State Model	-0.032 (0.027)	-0.029 (0.024)	-0.017 (0.028)	-0.018 (0.028)
Trial 6	Kaplan-Meier	0.037 (0.036)	0.026 (0.042)	0.048 (0.035)	0.031 (0.041)
	Full Multi-State Model	0.009 (0.030)	0.004 (0.032)	0.033 (0.029)	0.032 (0.038)
Trial 7	Kaplan-Meier	0.080 (0.035)	0.122 (0.045)	0.037 (0.035)	0.082 (0.041)
	Full Multi-State Model	0.039 (0.029)	0.051 (0.028)	0.040 (0.030)	0.043 (0.028)
Trial 8	Kaplan-Meier	0.105 (0.028)	0.112 (0.031)	0.084 (0.028)	0.116 (0.031)
	Full Multi-State Model	0.095 (0.025)	0.085 (0.024)	0.103 (0.026)	0.096 (0.025)
Trial 10	Kaplan-Meier	0.004 (0.019)	0.012 (0.021)	0.007 (0.018)	0.018 (0.022)
	Full Multi-State Model	0.012 (0.016)	0.015 (0.015)	0.010 (0.017)	0.008 (0.017)
Trial 11	Kaplan-Meier	-0.0001 (0.021)	0.026 (0.030)	-0.005 (0.021)	-0.033 (0.026)
	Full Multi-State Model	0.011 (0.018)	0.017 (0.017)	-0.004 (0.019)	-0.008 (0.019)
Trial 12	Kaplan-Meier	0.018 (0.017)	0.029 (0.019)	0.032 (0.017)	0.048 (0.020)
	Full Multi-State Model	0.017 (0.015)	0.008 (0.013)	0.033 (0.017)	0.035 (0.016)

$^*\Delta S(5) = P(T > 5|Z_i = 1) - P(T > 5|Z_i = 0)$

$^{**}\Delta DFS(3) = P(DFS > 3|Z_i = 1) - P(DFS > 3|Z_i = 0)$

Web Appendix C: Efficiency gains obtained through the proposed method utilizing the multi-state cure model: imputation based estimates.

Web Table 3: Treatment effect on survival from original data, reduced follow-up data and reduced follow-up data with imputation

Study	Data	Log-Rank P-Value	Cox model Log Hazard Ratio (SE)	5 year KM Estimate (SE)
1	Original	0.136	-0.28 (0.188)	0.090 (0.062)
	Reduced follow-up	0.035	-0.45 (0.214)	0.098 (0.070)
	Imputed Original	0.149	-0.31 (0.183)	0.092 (0.062)
	Imputed Reduced follow-up	0.047	-0.39 (0.197)	0.104 (0.063)
	Conlon, *et al.* (2011) method	0.117	-0.31 (0.199)	0.101 (0.065)
2	Original	0.097	-0.25 (0.155)	0.057 (0.049)
	Reduced follow-up	0.255	-0.20 (0.179)	0.039 (0.049)
	Imputed Original	0.105	-0.24 (0.154)	0.054 (0.048)
	Imputed Reduced follow-up	0.187	-0.23 (0.176)	0.051 (0.049)
	Conlon *et al.* (2011) method	0.203	-0.22 (0.175)	0.053 (0.050)
4	Original	0.719	0.06 (0.111)	-0.023 (0.037)
	Reduced follow-up	0.912	-0.005 (0.134)	0.027 (0.043)
	Imputed Original	0.704	0.05 (0.109)	-0.021 (0.037)
	Imputed Reduced follow-up	0.843	0.003 (0.132)	-0.005 (0.038)
	Conlon *et al.* (2011) method	0.739	0.05 (0.131)	-0.007 (0.038)
5	Original	0.355	0.09 (0.109)	-0.023 (0.031)
	Reduced follow-up	0.459	0.10 (0.134)	-0.009 (0.035)
	Imputed Original	0.464	0.06 (0.108)	-0.021 (0.031)
	Imputed Reduced follow-up	0.374	0.11 (0.129)	-0.020 (0.032)
	Conlon *et al.* (2011) method	0.443	0.09 (0.121)	-0.019 (0.034)
6	Original	0.695	-0.04 (0.101)	0.037 (0.036)
	Reduced follow-up	0.518	-0.08 (0.126)	0.026 (0.042)
	Imputed Original	0.734	-0.04 (0.100)	0.037 (0.036)
	Imputed Reduced follow-up	0.451	-0.09 (0.122)	0.024 (0.037)
	Conlon *et al.* (2011) method	0.578	-0.06 (0.119)	0.019 (0.036)
7	Original	0.053	-0.20 (0.115)	0.080 (0.035)
	Reduced follow-up	0.027	-0.33 (0.156)	0.122 (0.045)
	Imputed Original	0.070	-0.21 (0.114)	0.077 (0.035)
	Imputed Reduced follow-up	0.027	-0.31 (0.147)	0.079 (0.036)
	Conlon *et al.* (2011) method	0.014	-0.35 (0.146)	0.081 (0.037)
8	Original	0.0004	-0.36 (0.103)	0.105 (0.028)
	Reduced follow-up	0.0005	-0.41 (0.119)	0.112 (0.031)
	Imputed Original	0.0004	-0.35 (0.102)	0.105 (0.028)
	Imputed Reduced follow-up	0.0005	-0.40 (0.117)	0.103 (0.029)
	Conlon *et al.* (2011) method	0.0004	-0.41 (0.116)	0.103 (0.030)
10	Original	0.827	-0.02 (0.076)	0.004 (0.019)
	Reduced follow-up	0.398	-0.08 (0.092)	0.012 (0.021)
	Imputed Original	0.788	-0.02 (0.076)	0.004 (0.018)
	Imputed Reduced follow-up	0.502	-0.06 (0.091)	0.009 (0.019)
	Conlon*et al.* (2011) method	0.505	-0.06 (0.088)	0.010 (0.019)
11	Original	0.907	0.007 (0.096)	-0.0001 (0.021)
	Reduced follow-up	0.456	-0.08 (0.118)	0.026 (0.030)
	Imputed Original	0.930	0.02 (0.095)	-0.002 (0.021)
	Imputed Reduced follow-up	0.623	-0.05 (0.117)	0.005 (0.022)
	Conlon *et al.* (2011) method	0.622	-0.05 (0.111)	0.007 (0.022)
12	Original	0.083	-0.14 (0.086)	0.018 (0.017)
	Reduced follow-up	0.273	-0.09 (0.097)	0.029 (0.019)
	Imputed Original	0.080	-0.14 (0.086)	0.018 (0.017)
	Imputed Reduced follow-up	0.255	-0.09 (0.094)	0.019 (0.018)
	Conlon *et al.* (2011) method	0.230	-0.10 (0.091)	0.018 (0.018)

Web Appendix D: Simulation results under a misspecified model

Two additional simulations were performed to assess the robustness of the model and proposed methods to model misspecification. Data for each simulation was generated assuming a multi-state cure model with a lognormal distribution for each of the four transition times. Recurrence times and death times were simulated under this model to give "original data" with long follow up. These times were then censored at an earlier time to give "censored data". The multi-state cure model with a Weibull baseline hazard function was then fit to the data. Model based estimates of five year survival were then obtained from the model fit to the "censored data" and the imputation strategy was performed on the "censored data" to give the "imputed censored data". Two trial settings were explored, one with a treatment effect (Trial 1b) and one without a treatment effect (Trial 2b). For each setting, we generate 500 data sets, each with 500 subjects per treatment arm, 750 subjects with stage 3 disease, and a five year accrual period with eight years of additional follow-up to provide the "original data". The "censored data" is obtained by censoring these data sets two years after the last accrual to provide a maximum of seven years of follow-up time. The probability of being cured of disease was generated using $p_i = \frac{exp(\gamma_0+\gamma_1 Z_i+\gamma_2 S_i)}{1+exp(\gamma_0+\gamma_1 Z_i+\gamma_2 S_i)}$, where Z_i denotes treatment group and S_i denotes stage. Each of these covariates are centered at 0 so that Z_i is equal to -0.5 (0.5) for the control (treatment) group and S_i is equal to -0.75 (0.25) for stage 2 (stage 3) disease. We set $(\gamma_0, \gamma_1, \gamma_2) = (-0.2, 0.5, -1.0)$ in trial 1b and $(\gamma_0, \gamma_1, \gamma_2) = (-0.2, 0.0, -1.0)$ in trials 2b. For each transitions $1 \rightarrow 4$, $2 \rightarrow 3$ and $2 \rightarrow 4$, we generate a time from a lognormal distribution with $\mu_{kj} = \beta_{0_{kj}} + \beta_{trt_{kj}} Z_i + \beta_{st_{kj}} S_i$ and σ_{kj}, where μ_{kj} is the mean and σ_{kj} is the standard deviation of the log time for transition kj, respectively. For those who are cured of disease, we generate a death time with $\sigma_{14} = 2$, $\beta_{0_{14}} = 7$ and the treatment and stage effects set to 0. For those who are not cured we generate a recurrence time with $(\sigma_{23}, \beta_{0_{23}}, \beta_{trt_{23}}, \beta_{st_{23}}) = (1.5, 4.5, 0.7, -0.5)$ in trial

1b and $(\sigma_{23}, \beta_{0_{23}}, \beta_{trt_{23}}, \beta_{st_{23}}) = (1.5, 4.5, 0, -0.5)$ in trial 2b. We generate a death time for those who are not cured with $\sigma_{24} = 2$, $\beta_{0_{24}} = 6$ and the treatment and stage effects set to 0. If the death time for uncured subjects is less than the recurrence time, then a $2 \rightarrow 4$ transition is made at the death time and the recurrence is censored at the death time. If the recurrence time is less than the death time, then a $2 \rightarrow 3$ transition is made at that time. For those who recur, the time between their recurrence and death is generated with $(\sigma_{34}, \beta_{0_{34}}, \beta_{trt_{34}}, \beta_{st_{34}}\beta_{Tr}) = (0.75, 5, 0, -0.3, 0.1)$. The results in Web Table 4 demonstrate that smaller efficiency gains are obtained when the model is misspecified, but the Type I error rate is still maintained.

Web Table 4: Simulation results using the multi-state cure model with Weibull baseline hazards when data is generated from a lognormal distribution

Data		Misspecified model, Treatment Effect, 2 year censored					
	Size of Log-Rank	Cox model Log Hazard Ratio (SD)	Log Hazard Ratio $\bar{SE}$	Coverage	$\Delta S(5)$ Estimate (SD)	$\Delta S(5)$ $\bar{SE}$	Coverage
Original (max 13 year follow-up)	0.936	-0.23 (0.068)	0.068		0.062 (0.025)	0.025	
7 year follow-up	0.940	-0.24 (0.070)	0.070		0.062 (0.025)	0.025	
Censored (max 7 year follow-up)	0.958	-0.26 (0.071)	0.071		0.062 (0.028)	0.027	
Censored, model based					0.063 (0.025)	0.025	
Imputed Censored	0.956	-0.26 (0.072)	0.072		0.068 (0.026)	0.027	
		Misspecified model, No Treatment Effect, 2 year censored					
Original (max 13 year follow-up)	0.042	0.004 (0.067)	0.068	0.96	-0.002 (0.025)	0.025	0.94
7 year follow-up	0.056	0.000 (0.072)	0.069	0.94	0.001 (0.025)	0.025	0.95
Censored (max 7 year follow-up)	0.052	0.002 (0.070)	0.072	0.95	-0.002 (0.027)	0.027	0.94
Censored, model based					-0.001 (0.024)	0.025	0.95
Imputed Censored	0.054	0.002 (0.070)	0.072	0.95	-0.001 (0.025)	0.027	0.95

$\Delta S(5) = P(T > 5 \mid Z_i = 1) - P(T > 5 \mid Z_i = 0)$

Web Appendix E: Efficiency gains obtained through the proposed method utilizing the multi-state cure model with less informative prior distributions

Sensitivity of the proposed methods to the prior distributions was explored for the two trials considered. Imputation based estimates and model based estimates of overall survival using Normal(0, 5^2) priors on the log(λ)'s, gamma priors with mean 1 and standard deviation 1.6 on the ρ's, and Normal(0,2^2) on all of the covariate coefficients in the hazard models and in the logistic model were obtained. There is some sensitivity to the priors in the point estimates, but similar efficiency gains to those obtained under the more informative priors are obtained for both the model based estimates and the imputation procedure.

Web Table 5: Model based and imputation based estimates using the multi-state cure model with less informative priors for two colon cancer trials

		Model based estimates		Imputation based estimated		
		$\Delta S(5)^*$	$\Delta DFS(3)^{**}$	Log-Rank P-Value	Cox model Log Hazard Ratio (SE)	5 year KM Estimate (SE)
Trial 3	Full Follow-up	0.074 (0.031)	0.110 (0.031)	0.002	-0.31 (0.098)	0.074 (0.031)
	Reduced Follow-up	0.115 (0.080)	0.210 (0.086)	0.045	-0.27 (0.131)	0.115 (0.080)
	Multi-state Model Using Reduced Follow-up Data	0.096 (0.033)	0.128 (0.032)	0.011	-0.30 (0.119)	0.080 (0.032)
Trial 9	Full Follow-up	0.034 (0.021)	0.032 (0.021)	0.041	-0.16 (0.077)	0.034 (0.021)
	Reduced Follow-up	0.050 (0.026)	0.042 (0.022)	0.105	-0.14 (0.087)	0.050 (0.026)
	Multi-state Model Using Reduced Follow-up Data	0.042 (0.023)	0.049 (0.022)	0.082	-0.15 (0.086)	0.035 (0.021)

$^*\Delta S(5) = P(T > 5|Z_i = 1) - P(T > 5|Z_i = 0)$

$^{**}\Delta DFS(3) = P(DFS > 3|Z_i = 1) - P(DFS > 3|Z_i = 0)$

Web Appendix F: Efficiency gains obtained through the proposed method utilizing the multi-state cure model with restrictions on the treatment coefficients.

Li, *et al.* (2011) showed that when an intermediate variable captures even just a modest amount of the treatment effect on the final outcome, efficiency gains of the estimated treatment effect on the final outcome can be achieved by shrinking the treatment effect estimate in the conditional distribution of the final outcome given the intermediate variable and treatment toward 0. In our setting, it is plausible that much of the treatment effect is captured in the recurrence event by affecting the probability of being cured of disease and the time to recurrence. Therefore, one strategy to potentially improve efficiency gains in the estimation of the treatment effect on overall survival is to fit the multi-state cure model with strong prior assumptions placed on the treatment effects of some transition times. Specifically, the treatment effect on time to death for those who are cured ($1 \rightarrow 4$ transition) and the treatment effect on time to death for those who are not cured but without recurrence ($2 \rightarrow 4$ transition) are likely close to zero as the treatment may affect the probability of being cured, but after this most likely has little or no effect on the hazard of death from other causes if the person does not die from cancer. The treatment effect on time to death after recurrence ($3 \rightarrow 4$ transition) is also likely near zero, as patients often go off treatment or start new treatment regimens after a recurrence. We apply these restricted models to the 12 colon cancer trials, one model with the above mentioned treatment effects shrunk towards zero with the use of tighter prior distributions and another restricted model with these treatment effects forced to be zero. The results of applying the imputation procedure to these models are shown in Web Table 6. Small additional efficiency gains were obtained for some trials using these restricted models.

Web Table 6: Treatment effect on survival from original data, reduced follow-up data and reduced follow-up data with imputation and restrictions on treatment coefficients

Study	Data	Log-Rank P-Value	Cox model Log Hazard Ratio (SE)	5 year KM Estimate (SE)
1	Original	0.136	-0.28 (0.188)	0.090 (0.062)
	Reduced follow-up	0.035	-0.45 (0.214)	0.098 (0.070)
	Imputed Reduced follow-up, $\beta_{14}, \beta_{24}, \beta_{34}$ shrunk to 0	0.045	-0.39 (0.197)	0.105 (0.063)
	Imputed Reduced follow-up, $\beta_{14} = \beta_{24} = \beta_{34} = 0$	0.040	-0.40 (0.197)	0.105 (0.063)
2	Original	0.097	-0.25 (0.155)	0.057 (0.049)
	Reduced follow-up	0.255	-0.20 (0.179)	0.039 (0.049)
	Imputed Reduced follow-up, $\beta_{14}, \beta_{24}, \beta_{34}$ shrunk to 0	0.199	-0.23 (0.177)	0.051 (0.049)
	Imputed Reduced follow-up, $\beta_{14} = \beta_{24} = \beta_{34} = 0$	0.179	-0.24 (0.176)	0.051 (0.049)
3	Original	0.002	-0.31 (0.098)	0.074 (0.031)
	Reduced follow-up	0.045	-0.27 (0.131)	0.115 (0.080)
	Imputed Reduced follow-up, $\beta_{14}, \beta_{24}, \beta_{34}$ shrunk to 0	0.010	-0.31 (0.118)	0.082 (0.033)
	Imputed Reduced follow-up, $\beta_{14} = \beta_{24} = \beta_{34} = 0$	0.005	-0.33 (0.117)	0.092 (0.033)
4	Original	0.719	0.06 (0.111)	-0.023 (0.037)
	Reduced follow-up	0.912	-0.005 (0.134)	0.027 (0.043)
	Imputed Reduced follow-up, $\beta_{14}, \beta_{24}, \beta_{34}$ shrunk to 0	0.832	0.02 (0.130)	-0.006 (0.038)
	Imputed Reduced follow-up, $\beta_{14} = \beta_{24} = \beta_{34} = 0$	0.841	0.01 (0.131)	-0.004 (0.038)
5	Original	0.355	0.09 (0.109)	-0.023 (0.031)
	Reduced follow-up	0.459	0.10 (0.134)	-0.009 (0.035)
	Imputed Reduced follow-up, $\beta_{14}, \beta_{24}, \beta_{34}$ shrunk to 0	0.405	0.10 (0.125)	-0.019 (0.032)
	Imputed Reduced follow-up, $\beta_{14} = \beta_{24} = \beta_{34} = 0$	0.459	0.09 (0.124)	-0.017 (0.032)
6	Original	0.695	-0.04 (0.101)	0.037 (0.036)
	Reduced follow-up	0.518	-0.08 (0.126)	0.026 (0.042)
	Imputed Reduced follow-up, $\beta_{14}, \beta_{24}, \beta_{34}$ shrunk to 0	0.384	-0.11 (0.121)	0.026 (0.038)
	Imputed Reduced follow-up, $\beta_{14} = \beta_{24} = \beta_{34} = 0$	0.387	-0.10 (0.119)	0.026 (0.037)
7	Original	0.053	-0.20 (0.115)	0.080 (0.035)
	Reduced follow-up	0.027	-0.33 (0.156)	0.122 (0.045)
	Imputed Reduced follow-up, $\beta_{14}, \beta_{24}, \beta_{34}$ shrunk to 0	0.026	-0.31 (0.144)	0.078 (0.036)
	Imputed Reduced follow-up, $\beta_{14} = \beta_{24} = \beta_{34} = 0$	0.037	-0.28 (0.142)	0.074 (0.036)
8	Original	0.0004	-0.36 (0.103)	0.105 (0.028)
	Reduced follow-up	0.0005	-0.41 (0.119)	0.112 (0.031)
	Imputed Reduced follow-up, $\beta_{14}, \beta_{24}, \beta_{34}$ shrunk to 0	0.0005	-0.40 (0.115)	0.104 (0.029)
	Imputed Reduced follow-up, $\beta_{14} = \beta_{24} = \beta_{34} = 0$	0.0005	-0.40 (0.115)	0.103 (0.029)
9	Original	0.041	-0.16 (0.077)	0.034 (0.021)
	Reduced follow-up	0.105	-0.14 (0.087)	0.050 (0.026)
	Imputed Reduced follow-up, $\beta_{14} = \beta_{24} = \beta_{34} = 0$	0.071	-0.15 (0.085)	0.035 (0.021)
	Imputed Reduced follow-up, hierarchical model	0.082	-0.15 (0.086)	0.035 (0.021)
10	Original	0.827	-0.02 (0.076)	0.004 (0.019)
	Reduced follow-up	0.398	-0.08 (0.092)	0.012 (0.021)
	Imputed Reduced follow-up, $\beta_{14}, \beta_{24}, \beta_{34}$ shrunk to 0	0.511	-0.06 (0.089)	0.008 (0.019)
	Imputed Reduced follow-up, $\beta_{14} = \beta_{24} = \beta_{34} = 0$	0.585	-0.05 (0.088)	0.007 (0.019)
11	Original	0.907	0.007 (0.096)	-0.0001 (0.021)
	Reduced follow-up	0.456	-0.08 (0.118)	0.026 (0.030)
	Imputed Reduced follow-up, $\beta_{14}, \beta_{24}, \beta_{34}$ shrunk to 0	0.758	-0.01 (0.114)	0.003 (0.022)
	Imputed Reduced follow-up, $\beta_{14} = \beta_{24} = \beta_{34} = 0$	0.812	0.008 (0.113)	-0.0002 (0.022)
12	Original	0.083	-0.14 (0.086)	0.018 (0.017)
	Reduced follow-up	0.273	-0.09 (0.097)	0.029 (0.019)
	Imputed Reduced follow-up, $\beta_{14}, \beta_{24}, \beta_{34}$ shrunk to 0	0.194	-0.12 (0.092)	0.020 (0.018)
	Imputed Reduced follow-up, $\beta_{14} = \beta_{24} = \beta_{34} = 0$	0.140	-0.12 (0.091)	0.023 (0.018)

Web Appendix G: Efficiency gains obtained through the proposed method and a hierarchical model.

Another way to extend the use of the multi-state model and potentially improve upon the efficiency gains is to borrow information across trials by use of a hierarchical model. The original multi-state models fit to each individual trial provide evidence for common effects of some covariates on the probability of cure and transition rates. In particular, the coefficients associated with age and stage in all of the sub-models were quite similar. In addition, the coefficients associated with T_r in the $3 \rightarrow 4$ transition and the shape parameters of the Weibull models were similar across trials. We can therefore use a hierarchical model to borrow information across trials and shrink selected parameters towards common values. To illustrate this, we let $\rho_{skj} \sim N(\rho_{kj}, \sigma^2_{\rho kj})I(\rho_{skj} \geq 0)$, $\beta_{STskj} \sim N(\beta_{STkj}, \sigma^2_{\beta_{ST}kj})$, $\beta_{AGEskj} \sim N(\beta_{AGEkj}, \sigma^2_{\beta_{AGE}kj})$, $\beta_{T_rs34} \sim N(\beta_{T_r34}, \sigma^2_{\beta_{T_r}34})$, $\gamma_{STskj} \sim N(\gamma_{STkj}, \sigma^2_{\gamma_{ST}kj})$, and $\gamma_{AGEskj} \sim N(\gamma_{AGEkj}, \sigma^2_{\gamma_{AGE}kj})$, where $kj = \{12, 23, 24, 34\}$ corresponds to the transition and $s = 1, ..., 12$ represents the study number. We place Gamma hyper-priors with mean 1 and standard deviation 1 on ρ_{kj} and on $\sigma_{\rho kj}$, $\sigma_{\beta_{ST}kj}$, $\sigma_{\beta_{AGE}kj}$, $\sigma_{\beta_{T_r}kj}$, $\sigma_{\gamma_{ST}kj}$, and $\sigma_{\gamma_{AGE}kj}$ and $N(0, 2^2)$ hyper-priors on $\beta_{AGEkj}, \beta_{STkj}$, β_{T_rkj}, γ_{STkj}, and γ_{AGEkj}. The remaining parameters are independent across studies. For the reduced follow-up data, we fit the hierarchical model separately 12 times, each time with 1 trial artificially censored and the remaining 11 with their full follow-up data. The parameter estimates obtained from the hierarchical models can then be used in estimating five year OS and three year DFS. Web Table 7 provides log-rank p-values, Cox model log hazard ratio estimates and standard errors and Kaplan-Meier 5 year survival estimates and standard errors obtained from fitting the hierarchical model and applying the imputation procedure. There are not, in general, additional efficiency gains, likely due to the fact that these are all randomized trials and thus estimates for age and stage are likely to be at most weakly correlated with the estimate for treatment.

Web Table 7: Treatment effect on survival from original data, reduced follow-up data and reduced follow-up data with imputation from a hierarchical model

Study	Data	Log-Rank P-Value	Cox model Log Hazard Ratio (SE)	5 year KM Estimate (SE)
1	Original	0.136	-0.28 (0.188)	0.090 (0.062)
	Reduced follow-up	0.035	-0.45 (0.214)	0.098 (0.070)
	Imputed Reduced follow-up, hierarchical model	0.057	-0.38 (0.203)	0.102 (0.063)
2	Original	0.097	-0.25 (0.155)	0.057 (0.049)
	Reduced follow-up	0.255	-0.20 (0.179)	0.039 (0.049)
	Imputed Reduced follow-up, hierarchical model	0.215	-0.22 (0.178)	0.048 (0.049)
3	Original	0.002	-0.31 (0.098)	0.074 (0.031)
	Reduced follow-up	0.045	-0.27 (0.131)	0.115 (0.080)
	Imputed Reduced follow-up, hierarchical model	0.027	-0.27 (0.122)	0.072 (0.033)
4	Original	0.719	0.06 (0.111)	-0.023 (0.037)
	Reduced follow-up	0.912	-0.005 (0.134)	0.027 (0.043)
	Imputed Reduced follow-up, hierarchical model	0.823	0.01 (0.132)	-0.002 (0.038)
5	Original	0.355	0.09 (0.109)	-0.023 (0.031)
	Reduced follow-up	0.459	0.10 (0.134)	-0.009 (0.035)
	Imputed Reduced follow-up, hierarchical model	0.385	0.11 (0.128)	-0.019 (0.032)
6	Original	0.695	-0.04 (0.101)	0.037 (0.036)
	Reduced follow-up	0.518	-0.08 (0.126)	0.026 (0.042)
	Imputed Reduced follow-up, hierarchical model	0.465	-0.09 (0.123)	0.023 (0.038)
7	Original	0.053	-0.20 (0.115)	0.080 (0.035)
	Reduced follow-up	0.027	-0.33 (0.156)	0.122 (0.045)
	Imputed Reduced follow-up, hierarchical model	0.026	-0.31 (0.147)	0.077 (0.036)
8	Original	0.0004	-0.36 (0.103)	0.105 (0.028)
	Reduced follow-up	0.0005	-0.41 (0.119)	0.112 (0.031)
	Imputed Reduced follow-up, hierarchical model	0.0005	-0.40 (0.117)	0.103 (0.029)
9	Original	0.041	-0.16 (0.077)	0.034 (0.021)
	Reduced follow-up	0.105	-0.14 (0.087)	0.050 (0.026)
	Imputed Reduced follow-up, hierarchical model	0.082	-0.15 (0.086)	0.035 (0.021)
10	Original	0.827	-0.02 (0.076)	0.004 (0.019)
	Reduced follow-up	0.398	-0.08 (0.092)	0.012 (0.021)
	Imputed Reduced follow-up, hierarchical model	0.488	-0.07 (0.091)	0.009 (0.019)
11	Original	0.907	0.007 (0.096)	-0.0001 (0.021)
	Reduced follow-up	0.456	-0.08 (0.118)	0.026 (0.030)
	Imputed Reduced follow-up, hierarchical model	0.672	-0.03 (0.116)	0.007 (0.022)
12	Original	0.083	-0.14 (0.086)	0.018 (0.017)
	Reduced follow-up	0.273	-0.09 (0.097)	0.029 (0.019)
	Imputed Reduced follow-up, hierarchical model	0.226	-0.11 (0.095)	0.019 (0.017)

References

Li, Y., Taylor, J.M.G., and Little, R.J.A. (2011). A shrinkage approach for estimating a treatment effect using intermediate biomarker data in clinical trials. *Biometrics* **67**, 1434–1441.