***Changes from Pre-planned Statistical Methods***

This section describes in detail the statistical methods used to compute the patient-specific meta-analysis estimates of the risk of recurrence. The originally planned statistical methods are described in the Appendix. The only change from these methods was that estimates of the hazard ratio for 5FU effect were made separately by stage II versus III instead of overall.

***Studies***

The studies were designated as follows:

Study 1: CALGB 9581

Study 2: SUNRISE

Study 3: NSABP C-07

***Model Fitting***

The Cox proportional hazards regression models for each study had the following terms:

1. The RS result as a continuous measure, fit as a linear term.
2. Number of nodes examined (<12 vs. ≥12).
3. T-stage T4 vs. T3 or less.
4. MMR status (deficient vs. proficient/unknown).
5. (NSABP C-07 and SUNRISE only) Stage (II, IIIA/B or IIIC).
6. (NSABP C-07 only) Oxaliplatin + 5FU vs. placebo + 5FU

Studies 1 and 2 used stratified cohort sampling (8), so in the analysis of these studies each patient was weighted by the inverse sampling fraction in the patient's sampling stratum, and the covariance matrix of the regression parameter estimators was estimated using the method of Lin and Wei (9).

***Recurrence Risk Estimation***

The analysis requires an estimate of the hazard ratio for 5FU treatment added to surgery versus surgery alone. This hazard ratio was estimated using a meta-analysis of the original QUASAR study (2) and a pooled analysis of NSABP trials (12).

The log-rank observed-minus-expected ($O - E$) statistic, and its variance $V$, were reported for recurrence by stage in the QUASAR trial (11). From these quantities, the log-rank statistic $Z = (O - E)/\sqrt{V}$ can be computed. Since patients were allocated to treatment with 5FU or

observation with equal probability, the log hazard ratio can be estimated using the method of Schoenfeld (19) by $Z\sqrt{4/D}$, where $D$ is the total number of recurrence events. The variance of this estimate is consistently estimated by $4/D$. The results of this calculation are in the table below.

Log-Rank Statistics and Estimate of Log Hazard Ratio from the QUASAR trial

| | Events/Patients | | Log-rank Statistics | | Log Hazard Ratio (5FU:observation) | |
|---|---|---|---|---|---|---|
| Stage | 5FU | Observation | $O-E$ | $V$ | Estimate | Variance |
| II | 164 / 1073 | 194 / 1073 | −17.6 | 89.5 | −0.197 | 0.0112 |
| III | 58 / 131 | 68 /129 | −10.0 | 31.3 | −0.318 | 0.0317 |

Wilkinson *et al.* (12) provide estimates of the hazard ratio for recurrence for 5FU plus surgery versus surgery alone by stage based on a pooled analysis of NSABP trials. For stage II, the hazard ratio estimate is 0.70 with 95% confidence interval (0.54, 0.91); the log hazard ratio estimate is thus $\ln 0.70 = -0.35667$ with an estimated standard error of $(\ln 0.91 - \ln 0.54)/\{2\Phi(0.975)\} = 0.133134$. For stage III, the hazard ratio estimate is 0.61 with 95% confidence interval (0.52, 0.73), so the log hazard ratio estimate is $\ln 0.61 = -0.49430$ with standard error $(\ln 0.73 - \ln 0.52)/\{2\Phi(0.975)\} = 0.086536$.

Combining the QUASAR and Wilkinson log hazard ratio estimates in a meta-analysis using inverse-variance weighting gives a 5FU log hazard ratio estimate for stage II of $\hat{\lambda}_{5FU}^{(Stage\,II)} = -0.25852$ with standard error $\hat{\sigma}_{\lambda_{5FU}^{(Stage\,II)}} = 0.0828$ and for stage III, $\hat{\lambda}_{5FU}^{(Stage\,III)} = -0.461$ with standard error $\hat{\sigma}_{\lambda_{5FU}^{(Stage\,III)}} = 0.0778$. These estimates correspond to a hazard ratio for surgery and 5FU versus surgery alone for stage II of 0.772 with 95% confidence interval (0.657, 0.908) and for stage III hazard ratio 0.631 with 95% confidence interval (0.542, 0.735).

The risks of recurrence at 1, 3 and 5 years after surgery were estimated using patient-specific meta-analysis with special populations (10), integrated with the meta-analysis 5FU treatment

effect log hazard ratio. Here the special populations (not common to all studies) are Stage IIIA/B and IIIC patients and patients treated with oxaliplatin. The risk estimates were constructed as follows.

Define the vector of covariates as

$$\mathbf{z} = \left( RS, I_{\{<12 \text{ nodes ex.}\}}, I_{\{T4\}}, I_{\{MMRD\}}, I_{\{IIIA/B\}}, I_{\{IIIC\}}, I_{\{\text{oxali}\}} \right)^{\mathrm{T}}$$

and define

$$\mathbf{z}_0 = \left( RS, I_{\{<12 \text{ nodes ex.}\}}, I_{\{T4\}}, I_{\{MMRD\}}, 0, 0, 0 \right)^{\mathrm{T}},$$
$$\mathbf{z}_{IIIA/B} = \left( 0, 0, 0, 0, 1, 0, 0 \right)^{\mathrm{T}},$$
$$\mathbf{z}_{IIIC} = \left( 0, 0, 0, 0, 0, 1, 0 \right)^{\mathrm{T}}$$

and

$$\mathbf{z}_{\text{oxali}} = \left( 0, 0, 0, 0, 0, 0, 1 \right)^{\mathrm{T}}.$$

Since the overall recurrence risk has decreased since the late 1990's, we used the events from the latest-enrolling two studies (studies 2 and 3) to estimate the baseline cumulative hazard, with risk modification for individual presenting patients based on the regression parameters from each study. Since study 1 enrolled only stage II patients, the baseline for this study were estimated using only the stage II patients in studies 2 and 3. Similarly, since no patient in studies 1 or 2 was treated with oxaliplatin, the baselines for those studies estimated using study 3 were based on patients not treated with oxaliplatin. For each study $k = 1, 2, 3$, define $I_i^{(k,2)}$ as the indicator for whether study 2 patient $i$ is included in the baseline cumulative hazard estimator using study 2. Define the indicators $I_i^{(k,3)}$ similarly for baseline cumulative hazard estimators using study 3.

Let $\hat{\boldsymbol{\beta}}_k = \left( \hat{\beta}_1^{(k)}, \hat{\beta}_2^{(k)}, \ldots, \hat{\beta}_7^{(k)} \right)^{\mathrm{T}}$ and $\hat{\mathbf{V}}_k$ be the estimated proportional hazards regression parameter vector and its estimated covariance matrix for study $k = 1, 2, 3$. Set $\hat{\beta}_5^{(1)} = 0, \hat{\beta}_6^{(1)} = 0$ and $\hat{\beta}_7^{(1)} = 0$ and set $\hat{\beta}_7^{(2)} = 0,$ and set all corresponding elements of $\hat{\mathbf{V}}_k$ to 0. Let $\mathbf{z}_i^{(k)}$ be the observed covariate vector for patient $i = 1, 2, \ldots, n_k$ in study *k*. The Breslow-method estimator of the baseline cumulative hazard function at time *T* using the regression coefficients for study *k* and the events for study 2 is

$$\hat{\Lambda}_0^{(k,2)}(T) = \int_0^T \frac{I_i^{(k,2)} s_i^{(2)} \, \mathrm{d}\, N^{(2)}(t)}{\sum_{i=1}^{n_2} I_i^{(k,2)} s_i^{(2)} Y_i^{(2)}(t) \exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}} \mathbf{z}_i^{(2)}\right)},$$

where $N^{(2)}(t)$ is the event-counting process for study 2, $s_i^{(2)}$ is the stratified cohort sampling weight, and $Y_i^{(2)}(t)$ is the indicator for whether patient $i$ in study 2 is in the risk set at time *t*.

The variance of $\hat{\Lambda}_0^{(k,2)}(T)$ due to the event count is consistently estimated by

$$\widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,2)}(T)\right\} = \int_0^T \frac{I_i^{(k,2)} \left(s_i^{(2)}\right)^2 \mathrm{d}\, N^{(2)}(t)}{\left\{\sum_{i=1}^{n_3} I_i^{(k,2)} s_i^{(2)} Y_i^{(2)}(t) \exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}} \mathbf{z}_i^{(2)}\right)\right\}^2}.$$

Since all patients in study 3 received 5FU in addition to surgery, the baseline cumulative hazard estimator for study *k* using the events from study 3 is

$$\hat{\Lambda}_0^{(k,3)}(T) = \int_0^T \frac{I_i^{(k,3)} \, \mathrm{d}\, N^{(3)}(t)}{\sum_{i=1}^{n_3} I_i^{(k,3)} Y_i^{(3)}(t) \exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}} \mathbf{z}_i^{(3)} + \left(1 - I_{Stage\,III,i}\right) \hat{\lambda}_{5FU}^{(Stage\,II)} + I_{Stage\,III,i} \hat{\lambda}_{5FU}^{(Stage\,II)}\right)}.$$

The variance of $\hat{\Lambda}_0^{(k,3)}(T)$ due to the event count is consistently estimated by

$$\widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,3)}(T)\right\} = \int_0^T \frac{I_i^{(k,3)} \, \mathrm{d}\, N^{(3)}(t)}{\left\{\sum_{i=1}^{n_3} I_i^{(k,3)} Y_i^{(3)}(t) \exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}} \mathbf{z}_i^{(3)} + \left(1 - I_{Stage\,III,i}\right) \hat{\lambda}_{5FU}^{(Stage\,II)} + I_{Stage\,III,i} \hat{\lambda}_{5FU}^{(Stage\,II)}\right)\right\}^2}.$$

A fixed effects meta-analysis baseline cumulative hazard estimator combining studies 2 and 3 is

$$\hat{\Lambda}_0^{(k,2+3)}(T) = \omega_2^{(\Lambda_0)} \hat{\Lambda}_0^{(k,2)}(T) + \omega_3^{(\Lambda_0)} \hat{\Lambda}_0^{(k,3)}(T),$$

where

$$\omega_{k,2}^{(\Lambda_0)} = \frac{1 \Big/ \widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,2)}(T)\right\}}{1 \Big/ \widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,2)}(T)\right\} + 1 \Big/ \widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,3)}(T)\right\}}$$

and

$$\omega_{k,3}^{(\Lambda_0)} = \frac{1 \Big/ \widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,3}(T)\right\}}{1 \Big/ \widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,2)}(T)\right\} + 1 \Big/ \widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,3)}(T)\right\}}.$$

Let $I_{5FU}$ and $I_{\text{oxali}}$ be the indicators for whether the patient is to receive 5FU and oxaliplatin (with 5FU). For an individual patient with covariate vector $\mathbf{z}$, the estimated natural logarithm of the cumulative hazard at time $T$ for study $k = 1, 2, 3$ is

$$\hat{\rho}_k(T;\mathbf{z}) = \hat{\boldsymbol{\beta}}_k^{\mathrm{T}}\mathbf{z} + I_{5FU}\left\{\left(1 - I_{Stage\,III,i}\right)\hat{\lambda}_{5FU}^{(Stage\,II)} + I_{Stage\,III,i}\hat{\lambda}_{5FU}^{(Stage\,II)}\right\} + \ln\hat{\Lambda}_0^{(k,2+3)}(T).$$

Defining the gradient operator $\nabla_{\hat{\boldsymbol{\beta}}_k} = \left(\partial/\partial\beta_1^{(k)}, \partial/\partial\beta_2^{(k)}, \ldots, \partial/\partial\beta_7^{(k)}\right)^{\mathrm{T}}$, setting the elements for regression parameters that do not exist in each study to 0, we have

$$\nabla_{\hat{\boldsymbol{\beta}}_k}\hat{\rho}_k(T;\mathbf{z}) = \mathbf{z} + \frac{\nabla_{\hat{\boldsymbol{\beta}}_k}\hat{\Lambda}_0^{(k,2+3)}(T)}{\hat{\Lambda}_0^{(k,2+3)}(T)},$$

with $\nabla_{\hat{\boldsymbol{\beta}}_k}\hat{\Lambda}_0^{(k,2+3)}(T) = -\boldsymbol{\gamma}^{(k,2+3)}(T)$, where

$$\boldsymbol{\gamma}^{(k,2+3)}(T) = \omega_{k,2}^{(\Lambda_0)}\int_0^T \frac{\sum_{i=1}^{n_2} I_i^{(k,2)} s_i^{(2)} Y_i^{(2)}(t)\exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}}\mathbf{z}_i^{(2)}\right)\mathbf{z}_i^{(2)}}{\sum_{i=1}^{n_2} I_i^{(k,2)} s_i^{(2)} Y_i^{(3)}(t)\exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}}\mathbf{z}_i^{(2)}\right)}\mathrm{d}\hat{\Lambda}_0^{(k,2)}(t)$$
$$+ \omega_{k,3}^{(\Lambda_0)}\int_0^T \frac{\sum_{i=1}^{n_3} I_i^{(k,3)} Y_i^{(3)}(t)\exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}}\mathbf{z}_i^{(3)} + I_{Stage\,II,i}\hat{\lambda}_{5FU}^{(Stage\,II)} + I_{Stage\,III,i}\hat{\lambda}_{5FU}^{(Stage\,II)}\right)\mathbf{z}_i^{(3)}}{\sum_{i=1}^{n_3} I_i^{(k,3)} Y_i^{(3)}(t)\exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}}\mathbf{z}_i^{(3)} + I_{Stage\,II,i}\hat{\lambda}_{5FU}^{(Stage\,II)} + I_{Stage\,III,i}\hat{\lambda}_{5FU}^{(Stage\,II)}\right)}\mathrm{d}\hat{\Lambda}_0^{(k,3)}(t).$$

Also, we have

$$\frac{\partial\hat{\rho}_k(T;\mathbf{z})}{\partial\hat{\lambda}_{5FU}^{Stage\,II}} = I_{5FU}I_{Stage\,II} + \frac{\partial\hat{\Lambda}_0^{(k,2+3)}(T)}{\partial\hat{\lambda}_{5FU}^{Stage\,II}}\Bigg/\hat{\Lambda}_0^{(k,2+3)}(T),$$

with

$$\frac{\partial\hat{\Lambda}_0^{(k,2+3)}(T)}{\partial\hat{\lambda}_{5FU}^{Stage\,II}} =$$
$$-\omega_{k,3}^{(\Lambda_0)}\int_0^T \frac{\sum_{i=1}^{n_3} Y_i^{(3)}(t)\exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}}\mathbf{z}_i^{(3)} + I_{Stage\,II,i}\hat{\lambda}_{5FU}^{(Stage\,II)} + I_{Stage\,III,i}\hat{\lambda}_{5FU}^{(Stage\,II)}\right)I_{Stage\,II,i}}{\sum_{i=1}^{n_3} Y_i^{(3)}(t)\exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}}\mathbf{z}_i^{(3)} + I_{Stage\,II,i}\hat{\lambda}_{5FU}^{(Stage\,II)} + I_{Stage\,III,i}\hat{\lambda}_{5FU}^{(Stage\,II)}\right)}\mathrm{d}\hat{\Lambda}_0^{(k,3)}(t)$$

and, similarly,

$$\frac{\partial\hat{\rho}_k(T;\mathbf{z})}{\partial\hat{\lambda}_{5FU}^{Stage\,III}} = I_{5FU}I_{Stage\,III} + \frac{\partial\hat{\Lambda}_0^{(k,2+3)}(T)}{\partial\hat{\lambda}_{5FU}^{Stage\,III}}\Bigg/\hat{\Lambda}_0^{(k,2+3)}(T),$$

with

$$\frac{\partial \hat{\Lambda}_0^{(k,2+3)}(T)}{\partial \hat{\lambda}_{5FU}^{Stage\,III}} =$$
$$-\omega_{k,3}^{(\Lambda_0)} \int_0^T \frac{\sum_{i=1}^{n_3} Y_i^{(3)}(t) \exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}} \mathbf{z}_i^{(3)} + I_{Stage\,II,i} \hat{\lambda}_{5FU}^{(Stage\,II)} + I_{Stage\,III,i} \hat{\lambda}_{5FU}^{(Stage\,III)}\right) I_{Stage\,III,i}}{\sum_{i=1}^{n_3} Y_i^{(3)}(t) \exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}} \mathbf{z}_i^{(3)} + I_{Stage\,II,i} \hat{\lambda}_{5FU}^{(Stage\,II)} + I_{Stage\,III,i} \hat{\lambda}_{5FU}^{(Stage\,III)}\right)} \mathrm{d}\hat{\Lambda}_0^{(k,3)}(t).$$

Therefore, for $k = 1, 2, 3$, the variance of $\hat{\rho}_k(T;\mathbf{z})$ is consistently estimated by

$$\hat{\sigma}_k^2(T;\mathbf{z}) = \left(\nabla_{\hat{\boldsymbol{\beta}}_k} \hat{\rho}_k(T;\mathbf{z})\right)^{\mathrm{T}} \hat{\mathbf{V}}_k \left(\nabla_{\hat{\boldsymbol{\beta}}_k} \hat{\rho}_k(T;\mathbf{z})\right) + \left\{\frac{\partial \hat{\rho}_k(T;\mathbf{z})}{\partial \hat{\lambda}_{5FU}^{(Stage\,II)}}\right\}^2 \hat{\sigma}^2_{\hat{\lambda}_{5FU}^{(Stage\,II)}} + \left\{\frac{\partial \hat{\rho}_k(T;\mathbf{z})}{\partial \hat{\lambda}_{5FU}^{(Stage\,III)}}\right\}^2 \hat{\sigma}^2_{\hat{\lambda}_{5FU}^{(Stage\,III)}}$$
$$+ \frac{\left\{\omega_{k,2}^{(\Lambda_0)}\right\}^2 \widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,2)}(T)\right\} + \left\{\omega_{k,3}^{(\Lambda_0)}\right\}^2 \widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,3)}(T)\right\}}{\left\{\hat{\Lambda}_0^{(k,2+3)}(T)\right\}^2}.$$

Using the special population PSMA method (10) to adjust for the effects of Stage and oxaliplatin treatment, a fixed effects PSMA log cumulative hazard estimate for recurrence during the first *T* = 1, 3 or 5 years is

$$\hat{\rho}(T;\mathbf{z}) = \sum_{k=1}^{3} \omega_k(\mathbf{z}_0)\, \hat{\rho}_k\left(T;\mathbf{z}^{\dagger}\right),$$

where

$$\mathbf{z}^{\dagger} = \mathbf{z}_0 + \frac{I_{\{IIIA/B\}}\mathbf{z}_{IIIA/B} + I_{\{IIIC\}}\mathbf{z}_{IIIC}}{\omega_2(\mathbf{z}_0) + \omega_3(\mathbf{z}_0)} + \frac{I_{\mathrm{oxali}}\mathbf{z}_{\mathrm{oxali}}}{\omega_3(\mathbf{z}_0)}$$

and

$$\omega_k(\mathbf{z}_0) = \frac{w_k(5;\mathbf{z}_0)}{\sum_{j=1}^{3} w_j(5;\mathbf{z}_0)}$$

with

$$w_k\left(5;\mathbf{z}_0\right) = \frac{1}{\hat{\sigma}_k^2(5;\mathbf{z}_0)}.$$

The time is set at 5 years for the calculation of the weights so that the 1-, 3- and 5-year risk estimates are mutually consistent.

The gradient of $\hat{\rho}(T;\mathbf{z})$ with respect to $\hat{\boldsymbol{\beta}}_k^{\mathrm{T}}, k = 1, 2, 3,$ is

$$\nabla_{\hat{\boldsymbol{\beta}}_k}\hat{\rho}\left(T;\mathbf{z}\right)=\omega_k(\mathbf{z}_0)\nabla_{\hat{\boldsymbol{\beta}}_k}\hat{\rho}_k\left(T;\mathbf{z}^{\dagger}\right).$$

Letting $t_i^{(2)}$ denote the time to event or censoring and $N_i^{(2)}\left(t\right)$ denote the event-counting process for patient $i=1,2,\ldots,n_2$ in study 2, the partial derivative of $\hat{\Lambda}_0^{(k,2)}(T)$ with respect to $\mathrm{d}\,N_i^{(2)}\left(t_i^{(2)}\right)$ is

$$\frac{\partial\hat{\Lambda}_0^{(k,2)}(T)}{\partial\,\mathrm{d}\,N_i^{(2)}\left(t_i^{(2)}\right)}=I_{\left\{t_i^{(2)}\le T\right\}}\frac{I_i^{(k,2)}s_i^{(2)}\,\mathrm{d}\,N_i^{(2)}\left(t_i^{(2)}\right)}{\sum_{j=1}^{n_2}I_i^{(k,2)}s_i^{(2)}Y_j^{(2)}\left(t_i^{(2)}\right)\exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}}\mathbf{z}_j^{(2)}\right)}=I_{\left\{t_i^{(2)}\le T\right\}}\,\mathrm{d}\,\hat{\Lambda}_0^{(k,2)}\left(t_i^{(2)}\right),$$

where $I_{\left\{t_i^{(2)}\le T\right\}}$ is the indicator function for $t_i^{(2)}\le T,$ and $\mathrm{d}\,\hat{\Lambda}_0^{(k,2)}\left(t_i^{(2)}\right)$ is the increment in the baseline cumulative hazard estimate at time $t_i^{(2)}$. Therefore,

$$\frac{\partial\hat{\rho}(T;\mathbf{z})}{\partial\,\mathrm{d}\,N_i^{(2)}\left(t_i^{(2)}\right)}=\sum_{k=1}^{3}\omega_{k,2}^{(\Lambda_0)}\omega_k\left(\mathbf{z}_0\right)I_{\left\{t_i^{(2)}\le T\right\}}\frac{\mathrm{d}\,\hat{\Lambda}_0^{(k,2)}\left(t_i^{(2)}\right)}{\hat{\Lambda}_0^{(k,2)}(T)}.$$

Similarly, letting $t_i^{(3)}$ denote the time to event or censoring and $N_i^{(3)}\left(t\right)$ denote the event-counting process for patient $i=1,2,\ldots,n_3$ in study 3,

$$\frac{\partial\hat{\rho}(T;\mathbf{z})}{\partial\,\mathrm{d}\,N_i^{(3)}\left(t_i^{(3)}\right)}=\sum_{k=1}^{3}\omega_{k,3}^{(\Lambda_0)}\omega_k\left(\mathbf{z}_0\right)I_{\left\{t_i^{(3)}\le T\right\}}\frac{\mathrm{d}\,\hat{\Lambda}_0^{(k,3)}\left(t_i^{(3)}\right)}{\hat{\Lambda}_0^{(k,3)}(T)}.$$

Since the number and timing of events are asymptotically independent of the regression parameter estimates (20), assuming the three studies represent independent samples of patients and using methods similar to those in Therneau and Grambsch (21), the variance of $\hat{\rho}(T;\mathbf{z})$ is consistently estimated by

$$\begin{aligned}\widehat{\mathrm{Var}}\left\{\hat{\rho}(T;\mathbf{z})\right\}=&\sum_{k=1}^{3}\left(\omega_k(\mathbf{z}_0)\nabla_{\hat{\boldsymbol{\beta}}_k}\hat{\rho}_k\left(T;\mathbf{z}^{\dagger}\right)\right)^{\mathrm{T}}\hat{\mathbf{V}}_k\left(\omega_k(\mathbf{z}_0)\nabla_{\hat{\boldsymbol{\beta}}_k}\hat{\rho}_k\left(T;\mathbf{z}^{\dagger}\right)\right)\\&+\left\{\sum_{k=1}^{3}\omega_k(\mathbf{z}_0)\frac{\partial\hat{\rho}_k\left(\mathbf{z}^{\dagger}\right)}{\partial\hat{\lambda}_{5FU}^{(Stage\,II)}}\right\}^2\hat{\sigma}^2_{\lambda_{5FU}^{(Stage\,II)}}+\left\{\sum_{k=1}^{3}\omega_k(\mathbf{z}_0)\frac{\partial\hat{\rho}_k\left(\mathbf{z}^{\dagger}\right)}{\partial\hat{\lambda}_{5FU}^{(Stage\,III)}}\right\}^2\hat{\sigma}^2_{\lambda_{5FU}^{(Stage\,III)}}\\&+\sum_{i=1}^{n_2}\left\{\frac{\partial\hat{\rho}(T;\mathbf{z})}{\partial\,\mathrm{d}\,N_i^{(2)}\left(t_i^{(2)}\right)}\right\}^2\mathrm{d}\,N_i^{(2)}\left(t_i^{(2)}\right)\\&+\sum_{i=1}^{n_3}\left\{\frac{\partial\hat{\rho}(T;\mathbf{z})}{\partial\,\mathrm{d}\,N_i^{(3)}\left(t_i^{(3)}\right)}\right\}^2\mathrm{d}\,N_i^{(3)}\left(t_i^{(3)}\right).\end{aligned}$$

The third term in the sum above can be written

$$\sum_{i=1}^{n_2}\left\{\frac{\partial\hat{\rho}(T;\mathbf{z})}{\partial\,\mathrm{d}N_i^{(2)}\left(t_i^{(2)}\right)}\right\}^2 \mathrm{d}N_i^{(2)}\left(t_i^{(2)}\right)=\left\{\mathbf{c}_2(\mathbf{z}_0)\right\}^{\mathrm{T}}\mathbf{X}_2(T)\left\{\mathbf{c}_2(\mathbf{z}_0)\right\},$$

where $\mathbf{c}_2(\mathbf{z}_0)=\left(\omega_{1,2}^{(\Lambda_0)}\omega_1(\mathbf{z}_0),\omega_{2,2}^{(\Lambda_0)}\omega_2(\mathbf{z}_0),\omega_{3,2}^{(\Lambda_0)}\omega_3(\mathbf{z}_0)\right)^{\mathrm{T}}$ and $\mathbf{X}_2(T)$ is the matrix with element $x_{kl}^{(2)}(T)$ in row *k* and column *l*, with

$$x_{kl}^{(2)}(T)=\int_0^T\frac{\mathrm{d}\hat{\Lambda}_0^{(k,2)}(t)}{\hat{\Lambda}_0^{(k,2)}(T)}\frac{\mathrm{d}\hat{\Lambda}_0^{(l,2)}(t)}{\hat{\Lambda}_0^{(l,2)}(T)}\mathrm{d}N^{(2)}(t).$$

Similarly, the fourth term in (2) can be written

$$\sum_{i=1}^{n_3}\left\{\frac{\partial\hat{\rho}(T;\mathbf{z})}{\partial\,\mathrm{d}N_i^{(3)}\left(t_i^{(3)}\right)}\right\}^2 \mathrm{d}N_i^{(3)}\left(t_i^{(3)}\right)=\left\{\mathbf{c}_3(\mathbf{z}_0)\right\}^{\mathrm{T}}\mathbf{X}_3(T)\left\{\mathbf{c}_3(\mathbf{z}_0)\right\},$$

where $\mathbf{c}_3(\mathbf{z}_0)=\left(\omega_{1,3}^{(\Lambda_0)}\omega_1(\mathbf{z}_0),\omega_{2,3}^{(\Lambda_0)}\omega_2(\mathbf{z}_0),\omega_{3,3}^{(\Lambda_0)}\omega_3(\mathbf{z}_0)\right)^{\mathrm{T}}$ and $\mathbf{X}_3(T)$ is matrix with element $x_{kl}^{(3)}(T)$ in row *k* and column *l*, with

$$x_{kl}^{(3)}(T)=\int_0^T\frac{\mathrm{d}\hat{\Lambda}_0^{(k,3)}(t)}{\hat{\Lambda}_0^{(k,3)}(T)}\frac{\mathrm{d}\hat{\Lambda}_0^{(l,3)}(t)}{\hat{\Lambda}_0^{(l,3)}(T)}\mathrm{d}N^{(3)}(t).$$

Transforming to the risk scale, the estimated risk of a recurrence by time *T* is $\hat{r}(T;\mathbf{z})=1-\exp\left[-\exp\left\{\hat{\rho}(T;\mathbf{z})\right\}\right]$. A level-$\alpha$ confidence interval for the recurrence risk in the first *T* years after surgery has endpoints $1-\exp\left(-\exp\left[\hat{\rho}(T;\mathbf{z})\pm\Phi^{-1}(1-\alpha/2)\sqrt{\widehat{\mathrm{Var}}\left\{\hat{\rho}(T;\mathbf{z})\right\}}\right]\right)$, where $\Phi$ denotes the cumulative distribution function of the standard normal distribution.

***Testing Analysis Assumptions***

The assumption that there is no interaction among the 5 prognostic factors (RS result, Stage, T-stage, number of nodes examined and MMR proficiency) was tested using meta-analysis Wald tests. For each of the 4 categorical factors, a model was fit to each study allowing different regression parameters across the levels of the categorical factor. A fixed-effect meta-analysis estimate for the difference(s) across levels in these parameters was constructed using inverse

variance weighting and the statistical significance of the resulting meta-analysis difference estimates was assessed using a Wald test at a significance level of 0.10.

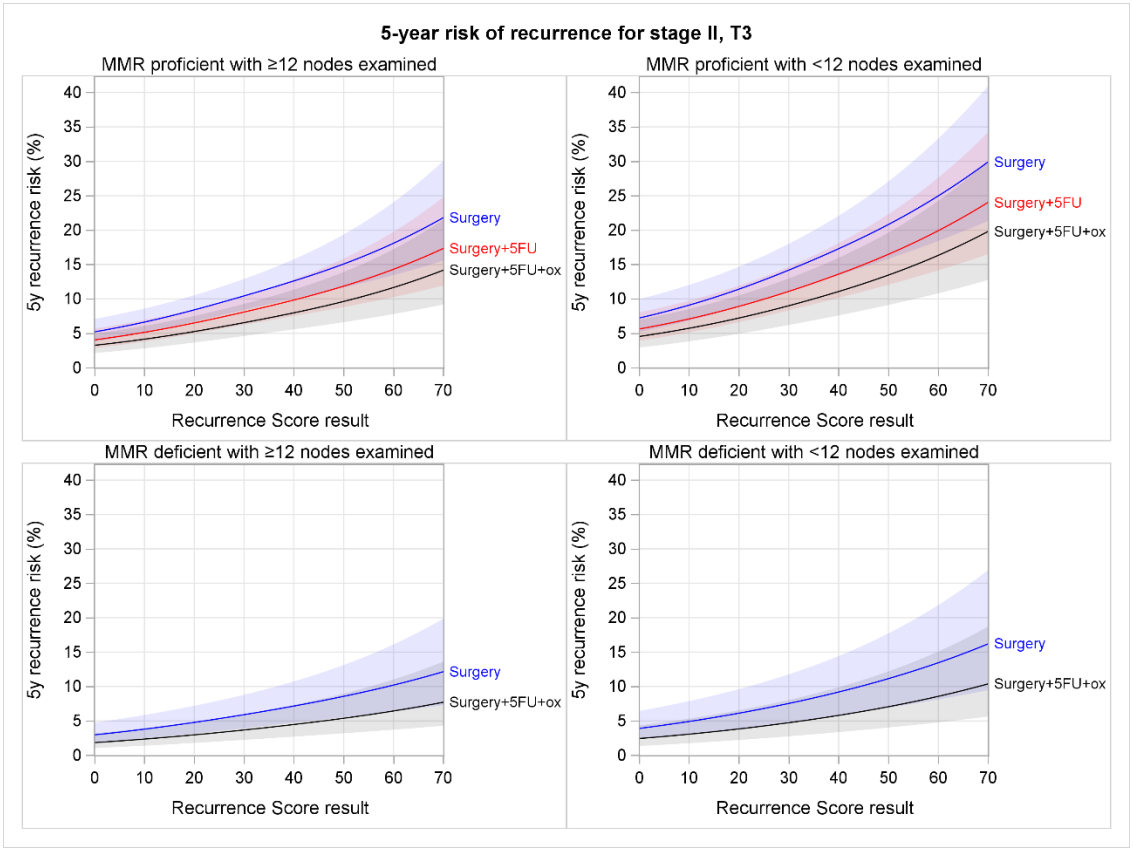
### *Additional References*

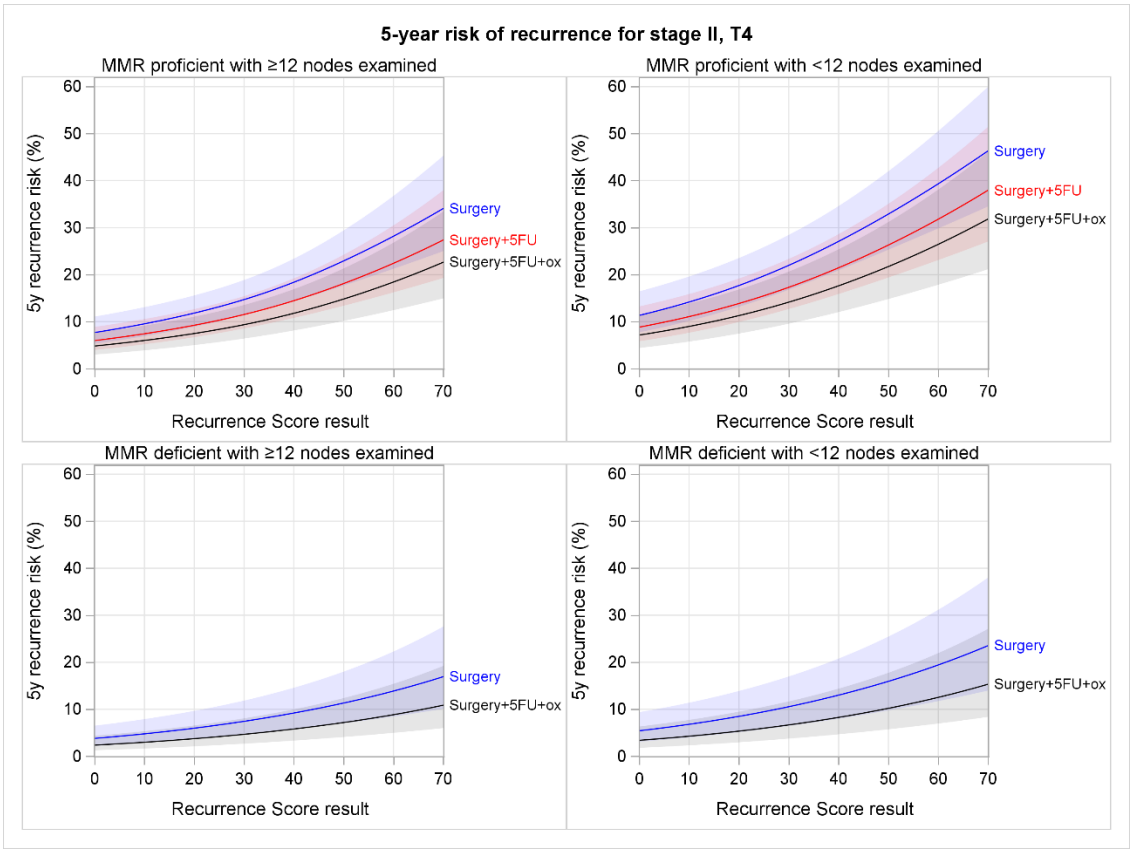
19. Schoenfeld D. The asymptotic properties of nonparametric tests for comparing survival distributions. Biometrika 1981;68:316-9.

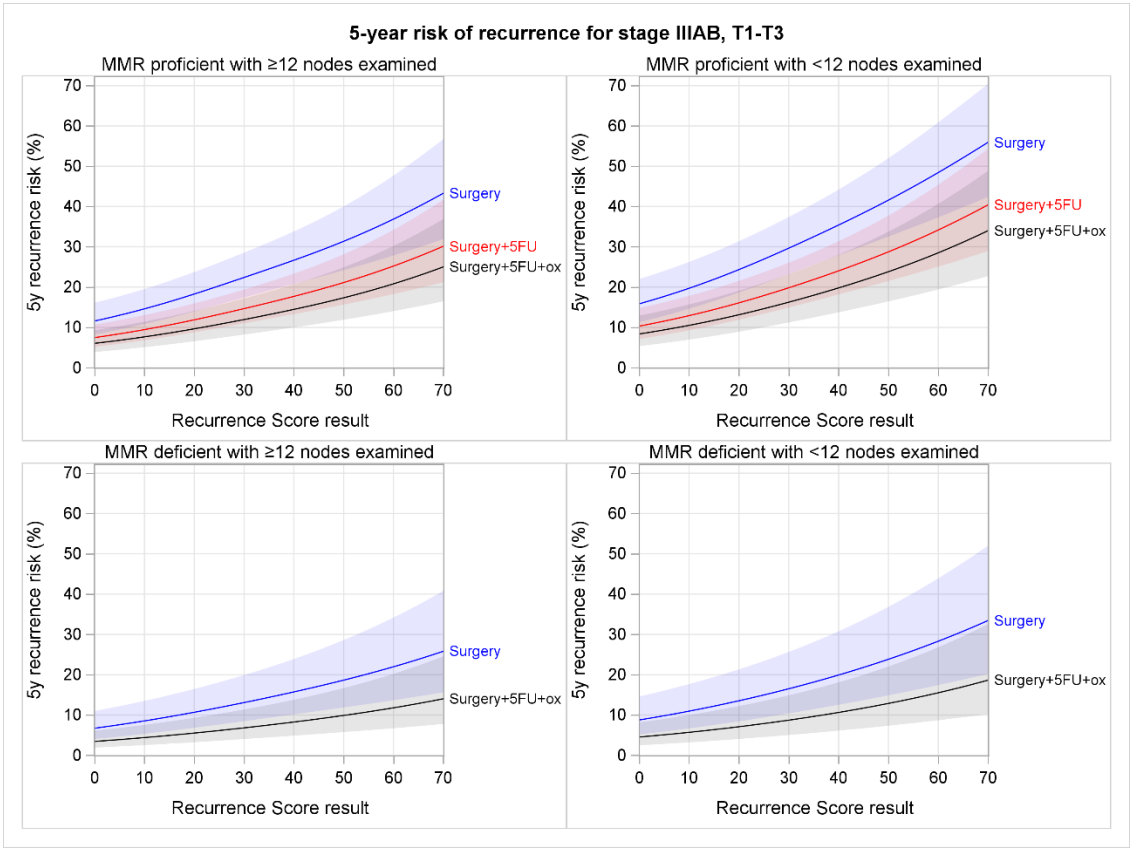
20. Tsiatis A. A large sample study of the estimates for the integrated hazard function in Cox's regression model for survival data. Annal Stat 1981;9:93-108.

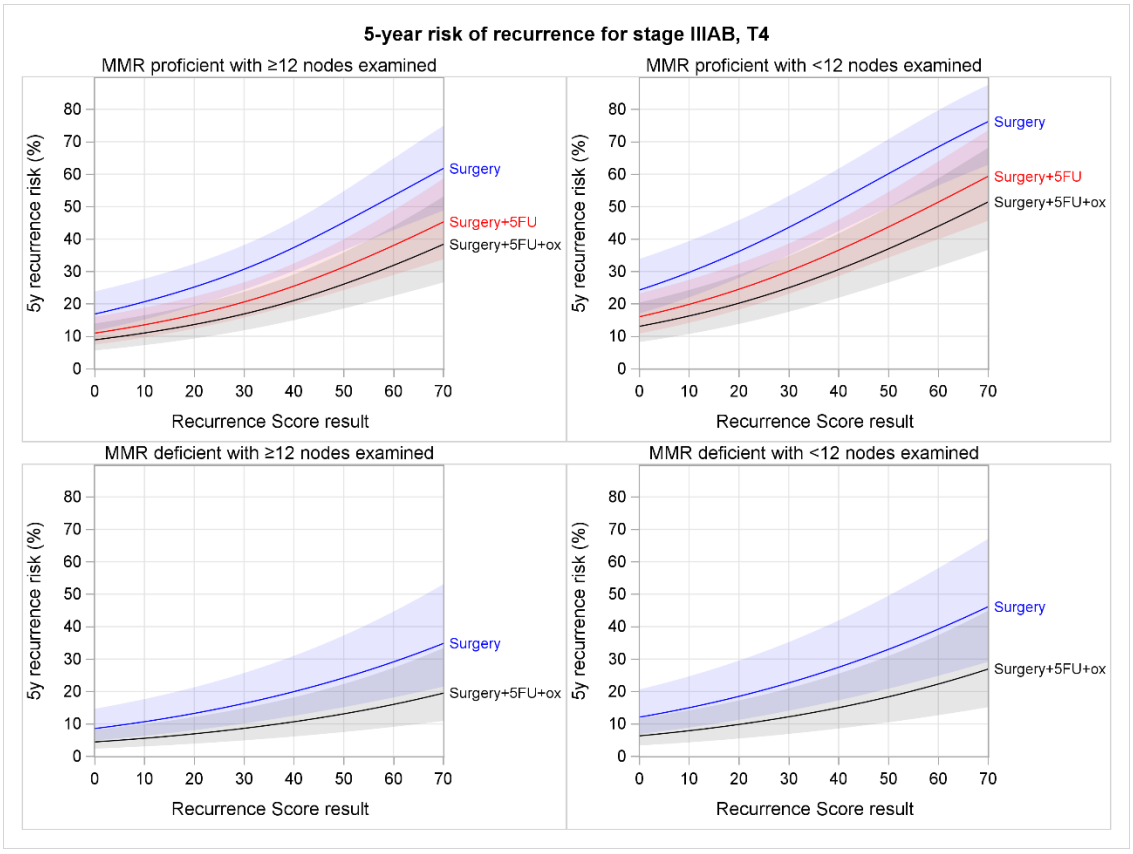
21. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. New York, NY, USA: Springer, 2020.

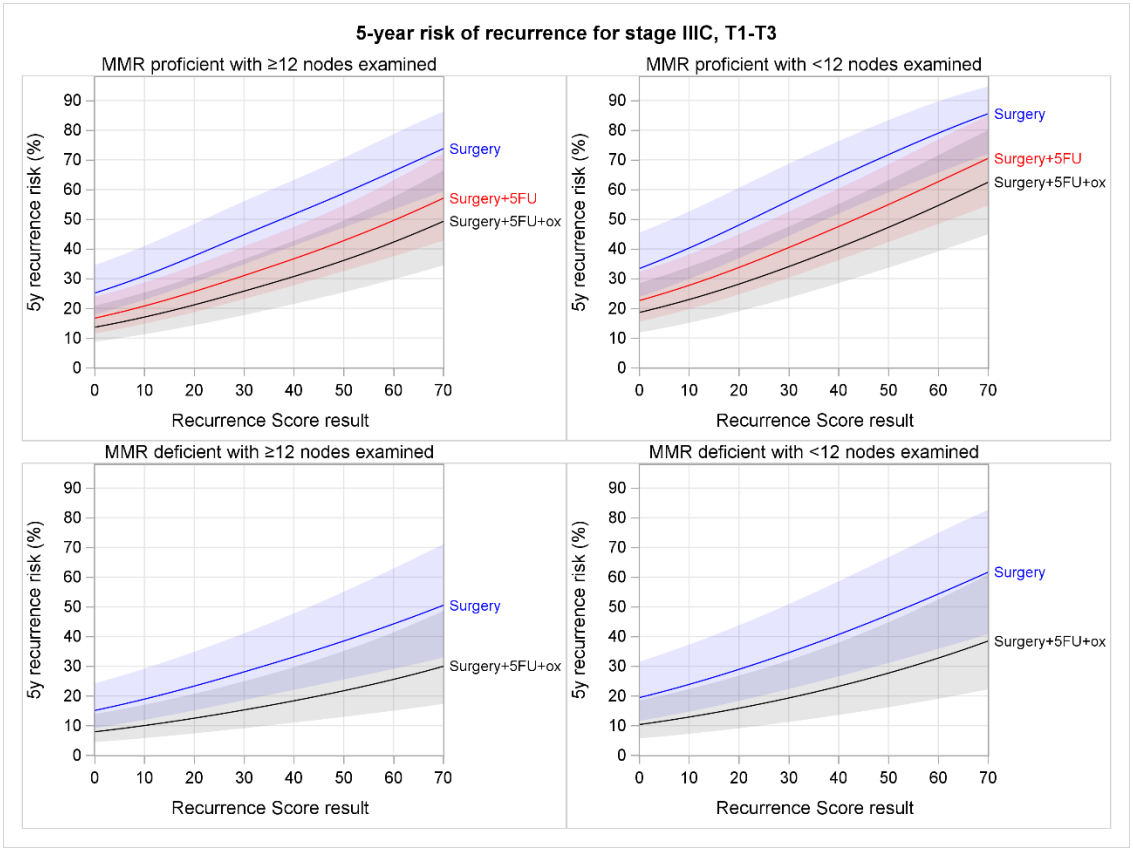
# 5-year Recurrence Risk Estimates

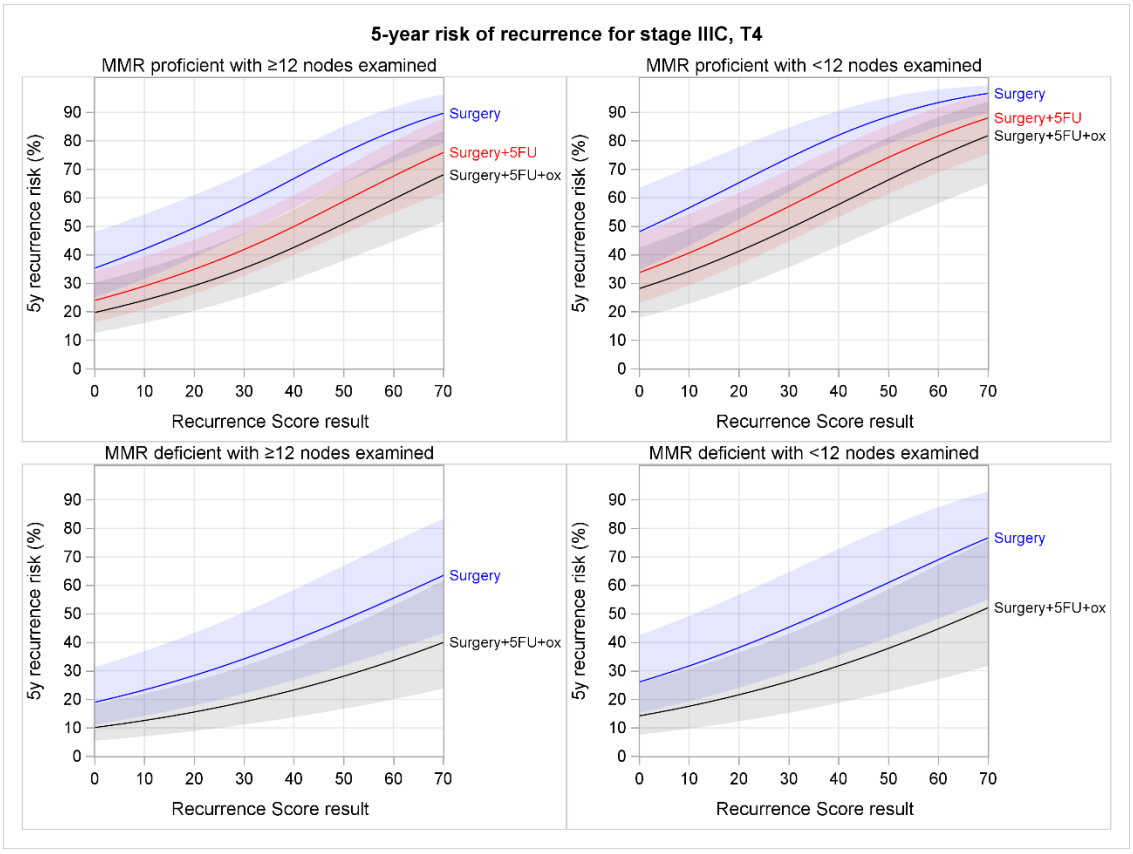

5-year risk of recurrence for stage II, T3
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
5y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox


5-year risk of recurrence for stage II, T4
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
5y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox

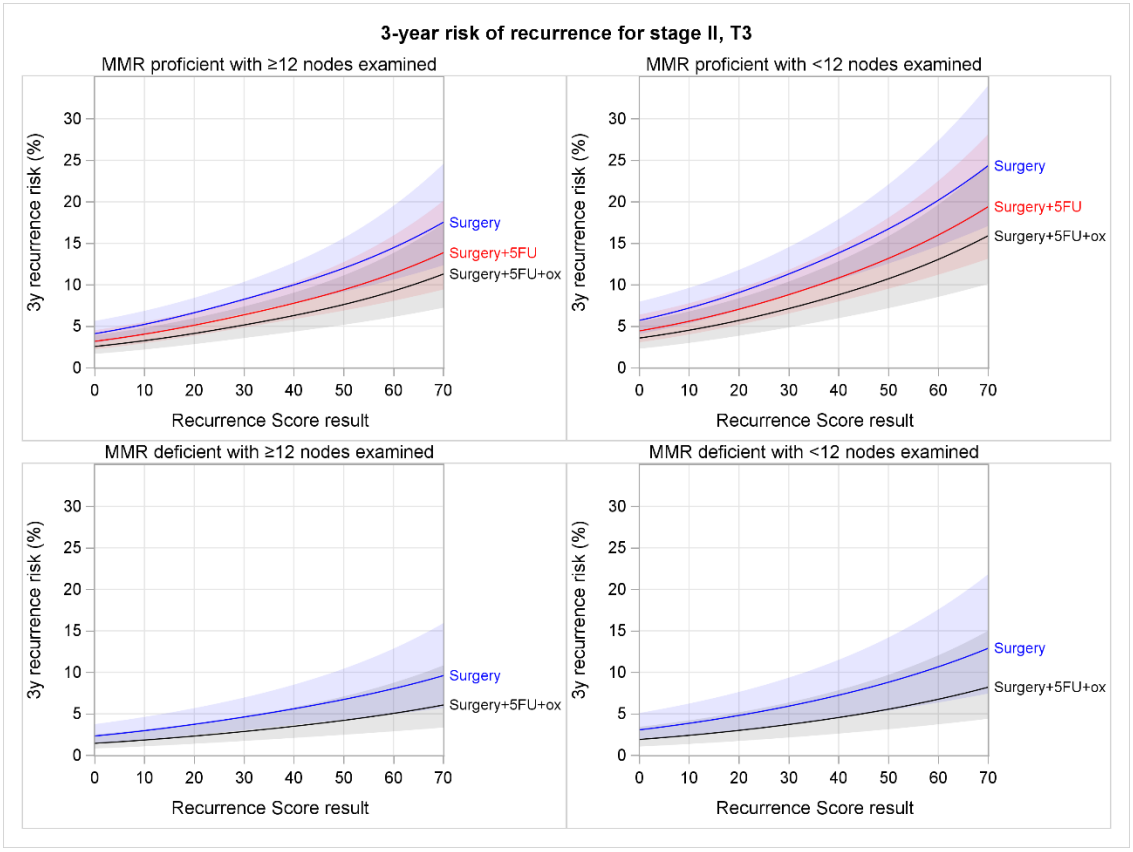

5-year risk of recurrence for stage IIIAB, T1-T3
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
5y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox

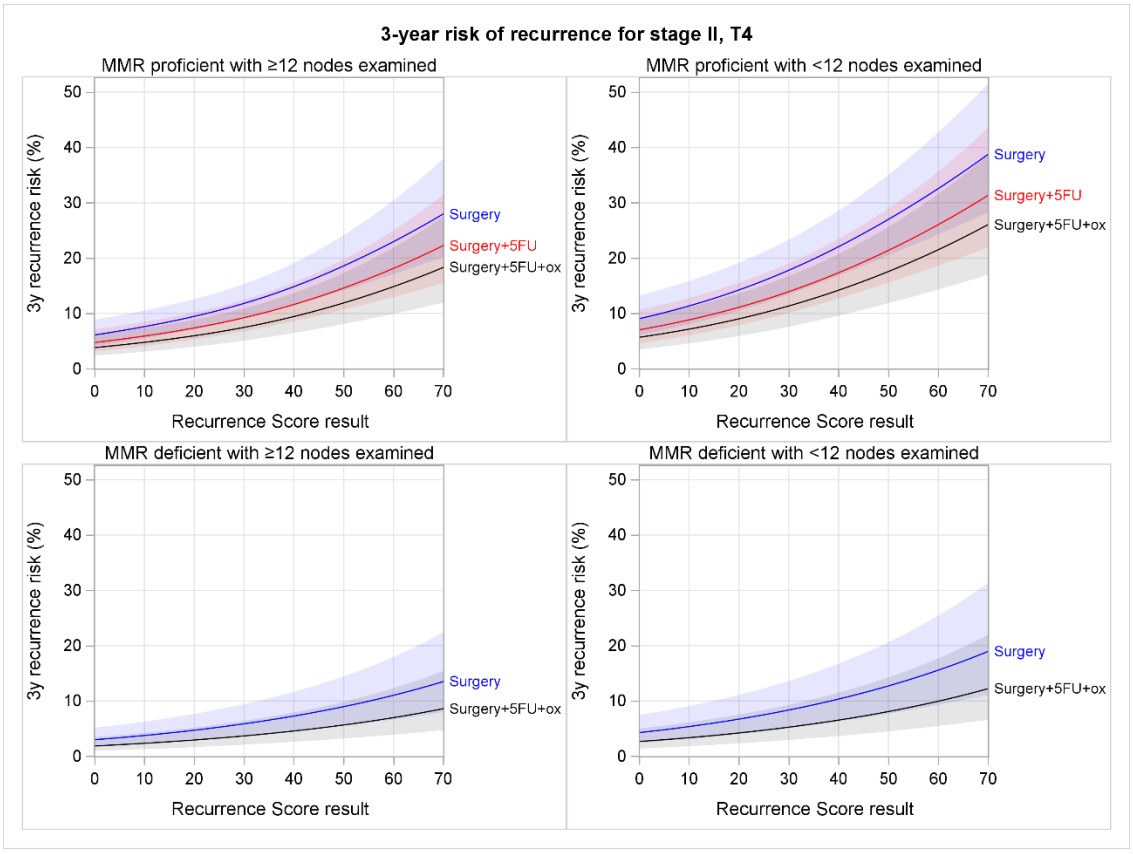

5-year risk of recurrence for stage IIIAB, T4
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
5y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox

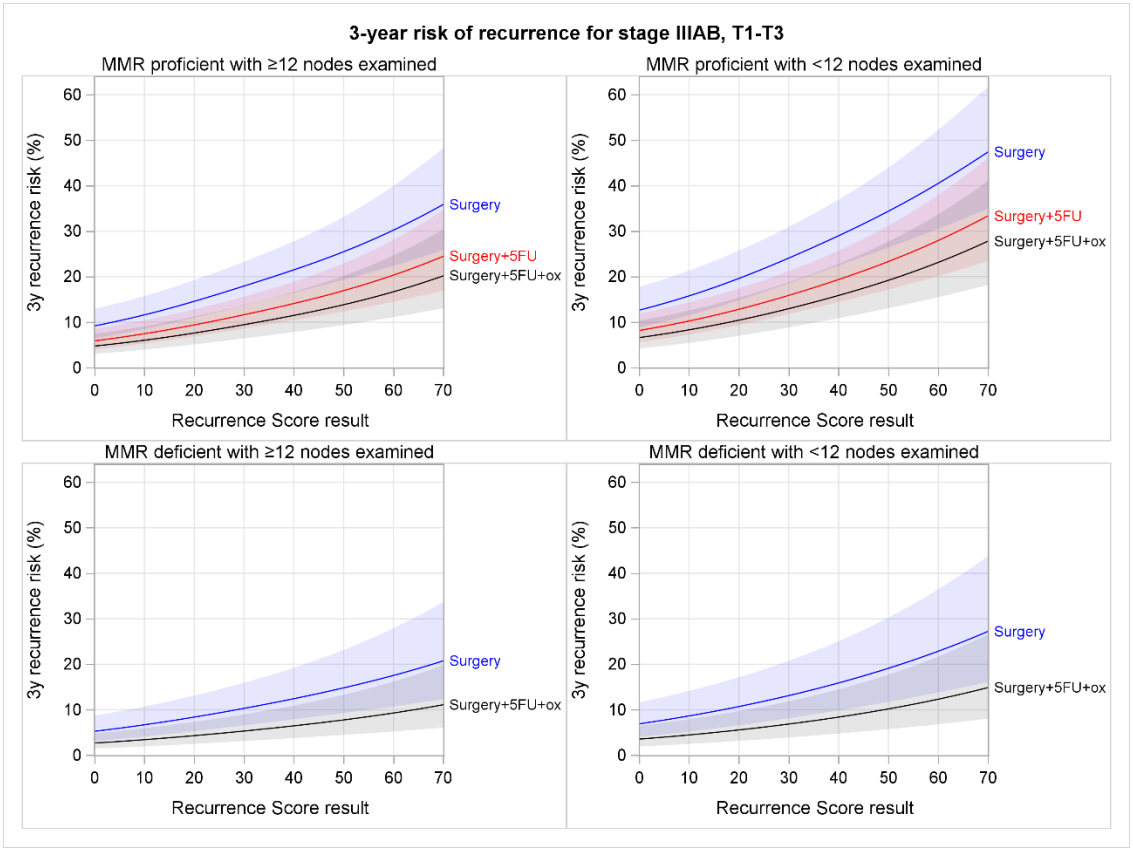

5-year risk of recurrence for stage IIIC, T1-T3
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
5y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox

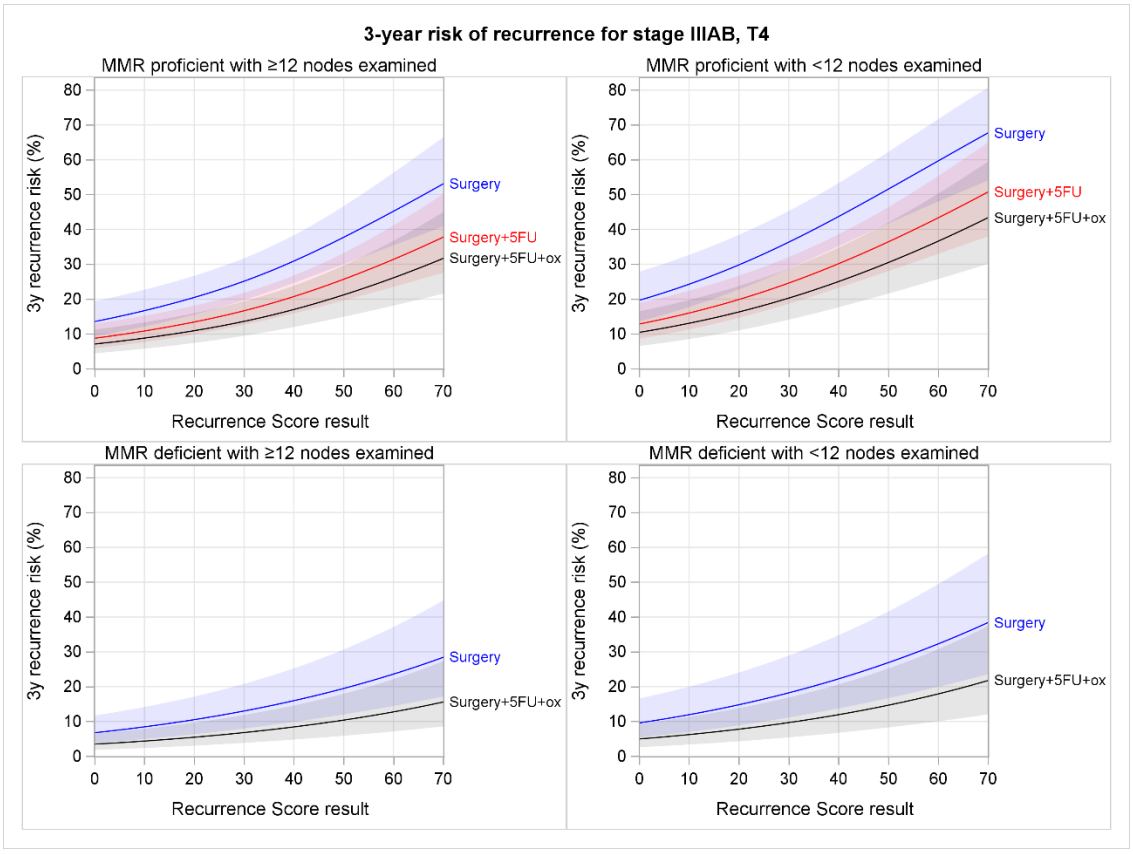

5-year risk of recurrence for stage IIIC, T4
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
5y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox

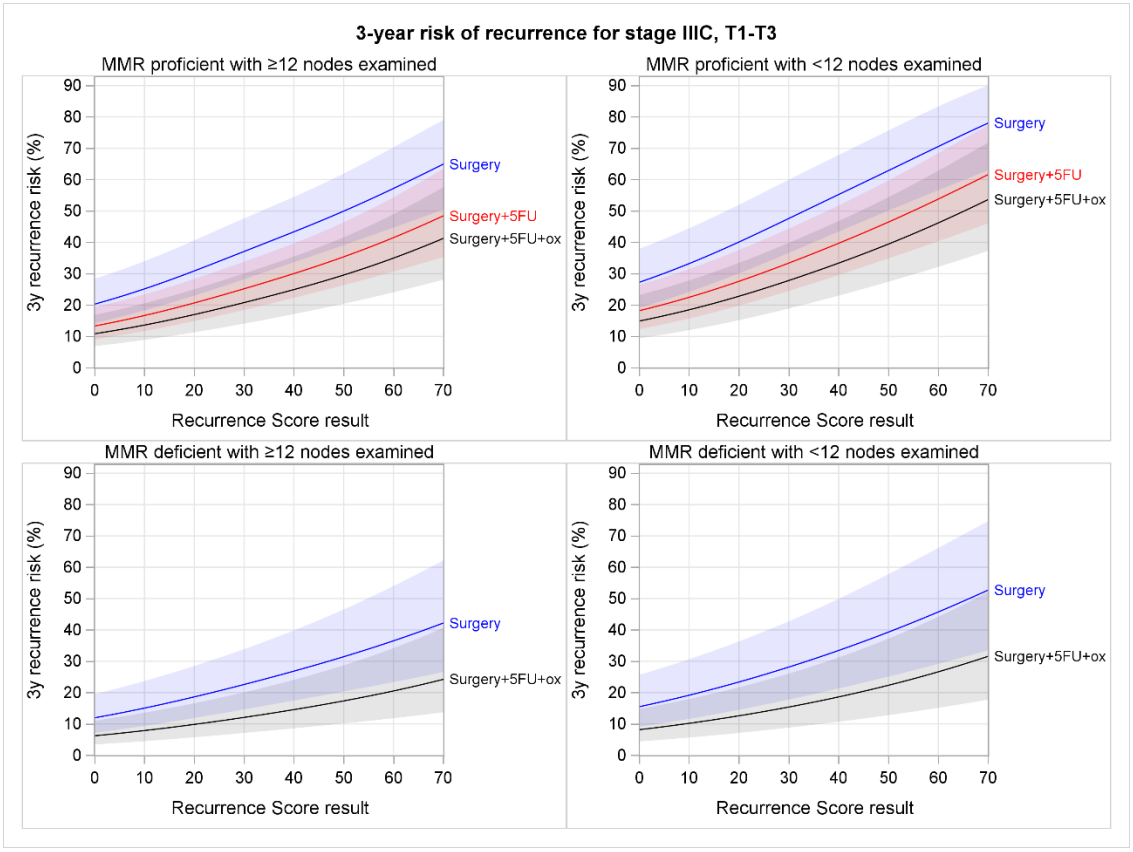
# 3-year Recurrence Risk Estimates

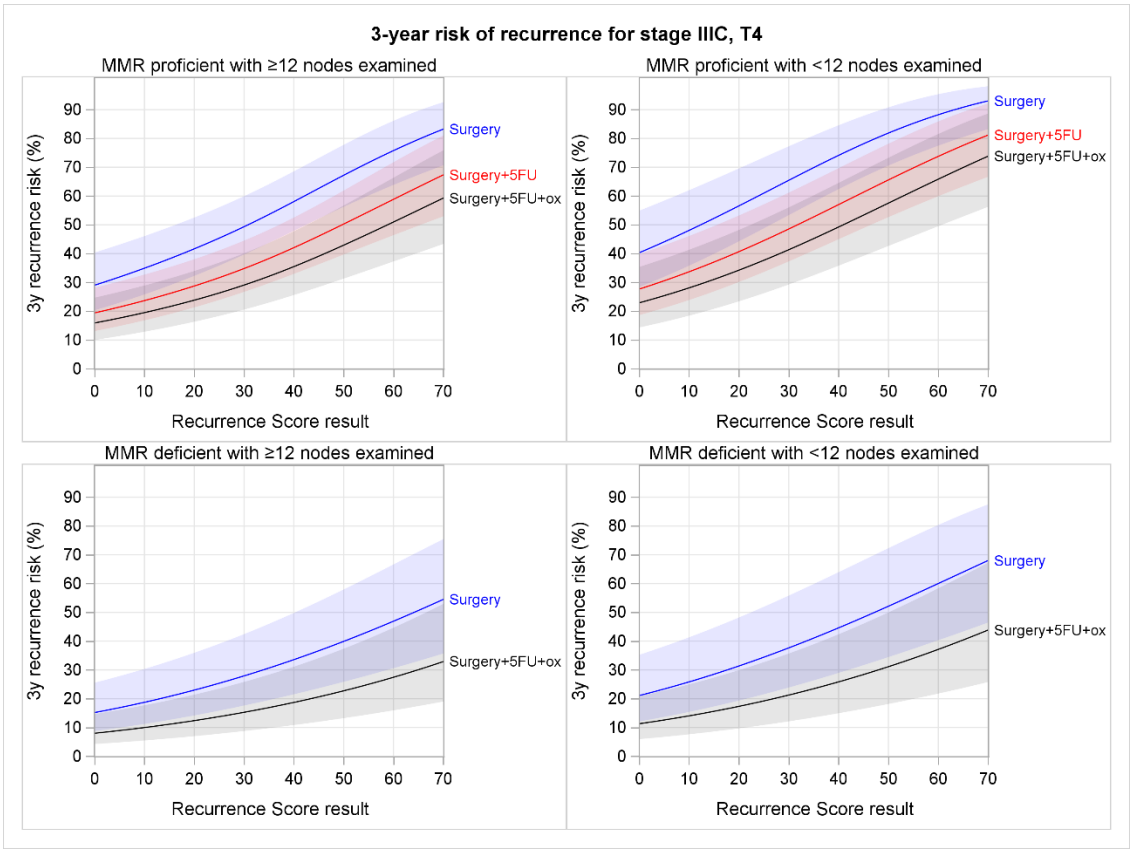

3-year risk of recurrence for stage II, T3
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
3y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox


3-year risk of recurrence for stage II, T4
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
3y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox

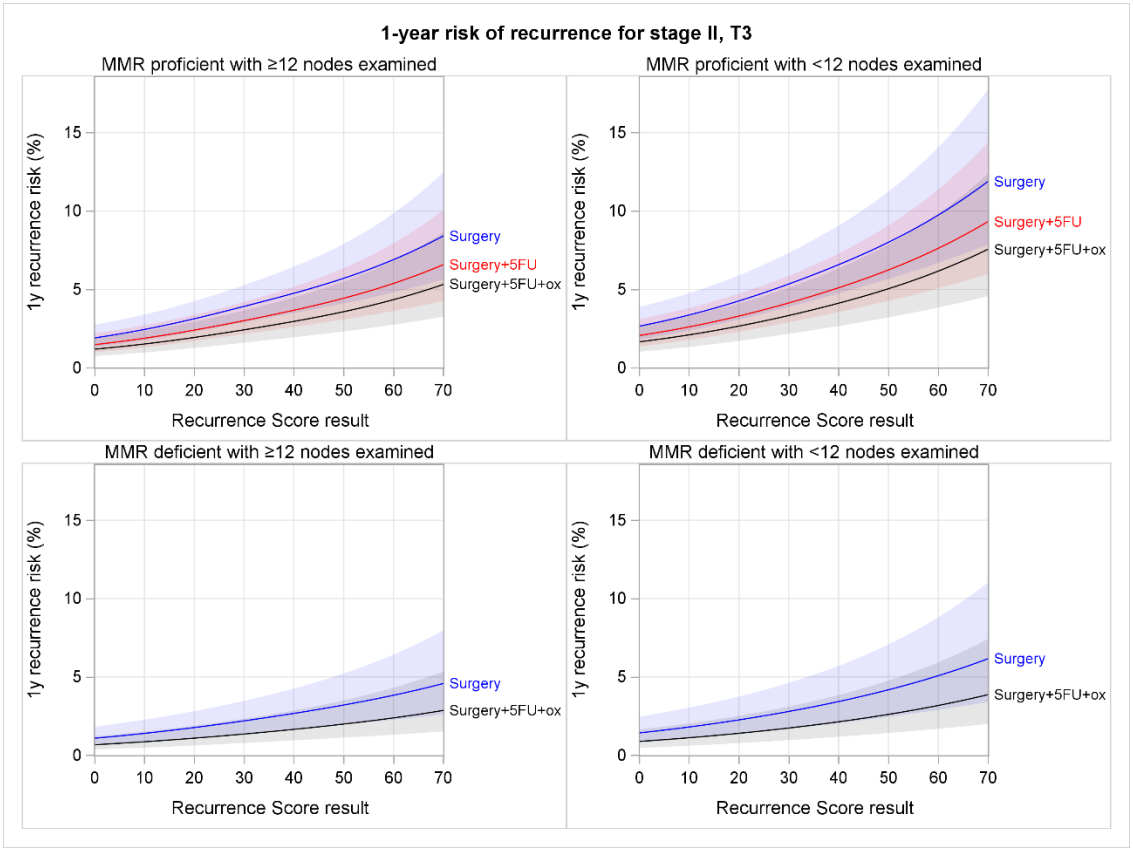

3-year risk of recurrence for stage IIIAB, T1-T3
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
3y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox

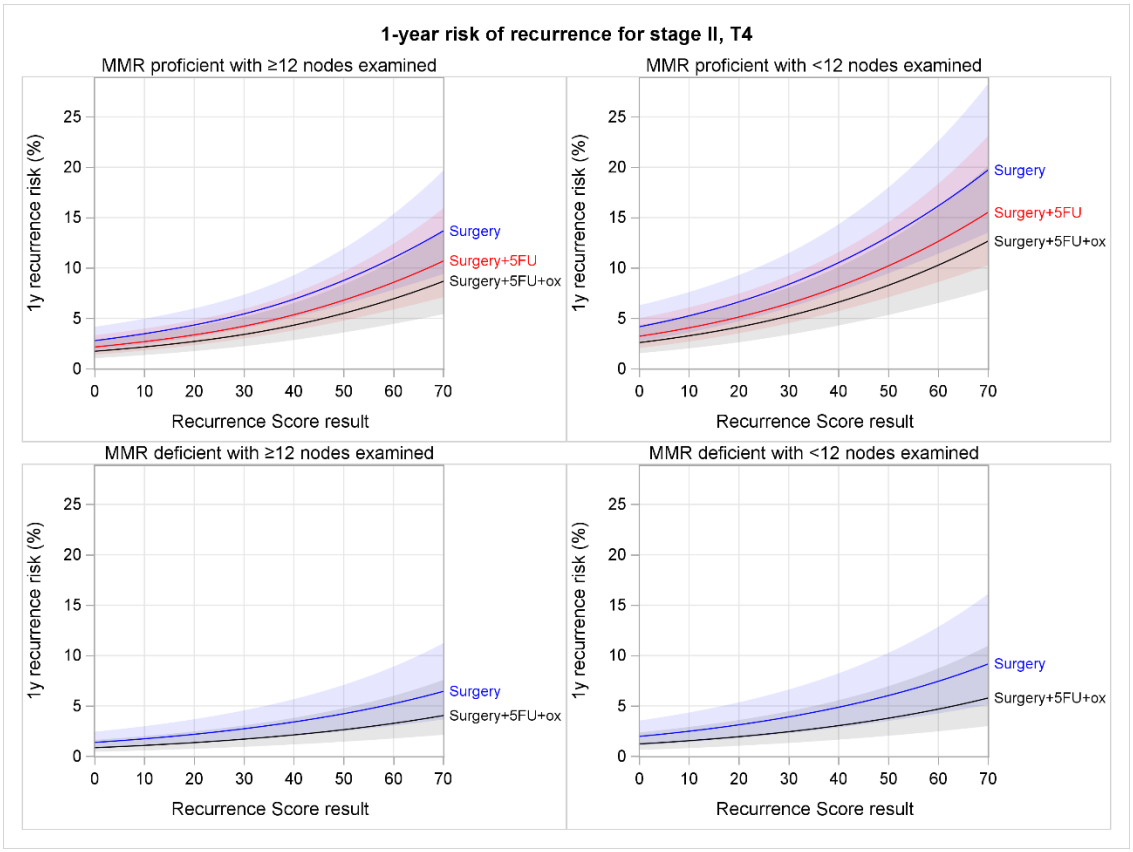

3-year risk of recurrence for stage IIIAB, T4
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
3y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox

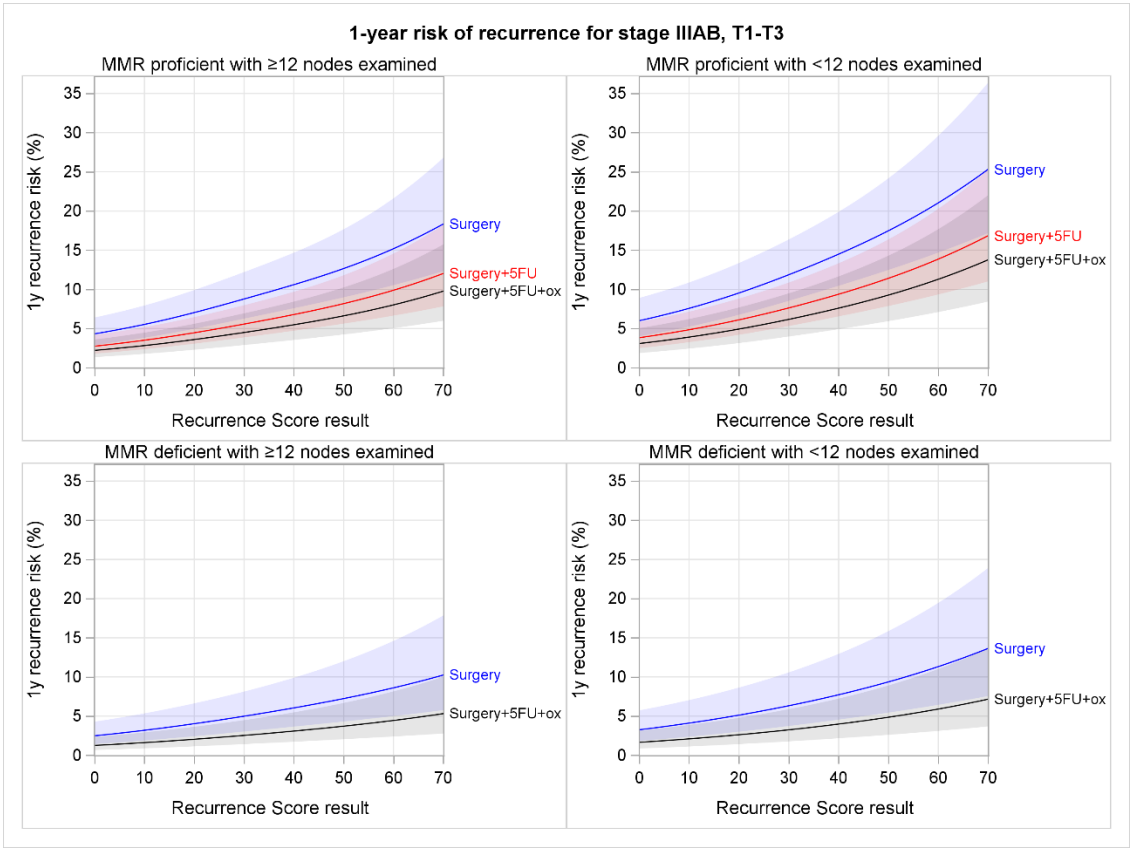

3-year risk of recurrence for stage IIIC, T1-T3
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
3y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox

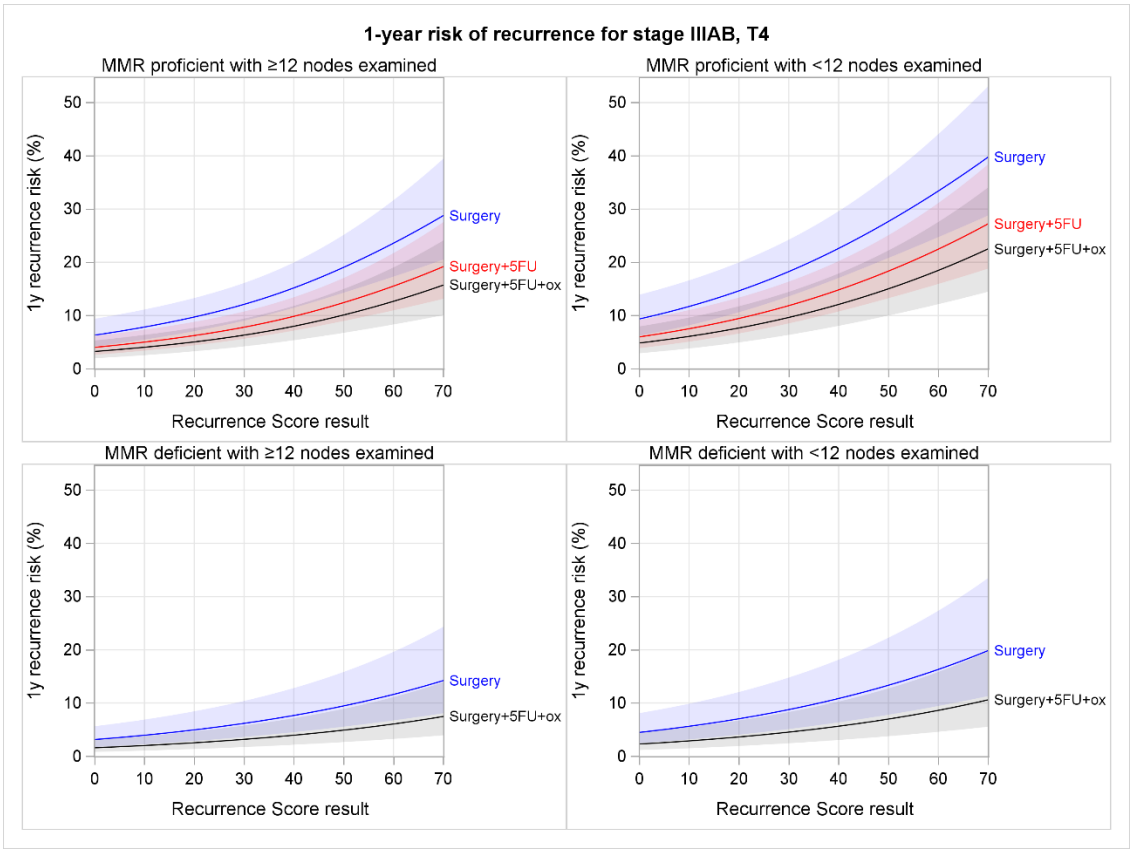

3-year risk of recurrence for stage IIIC, T4
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
3y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox

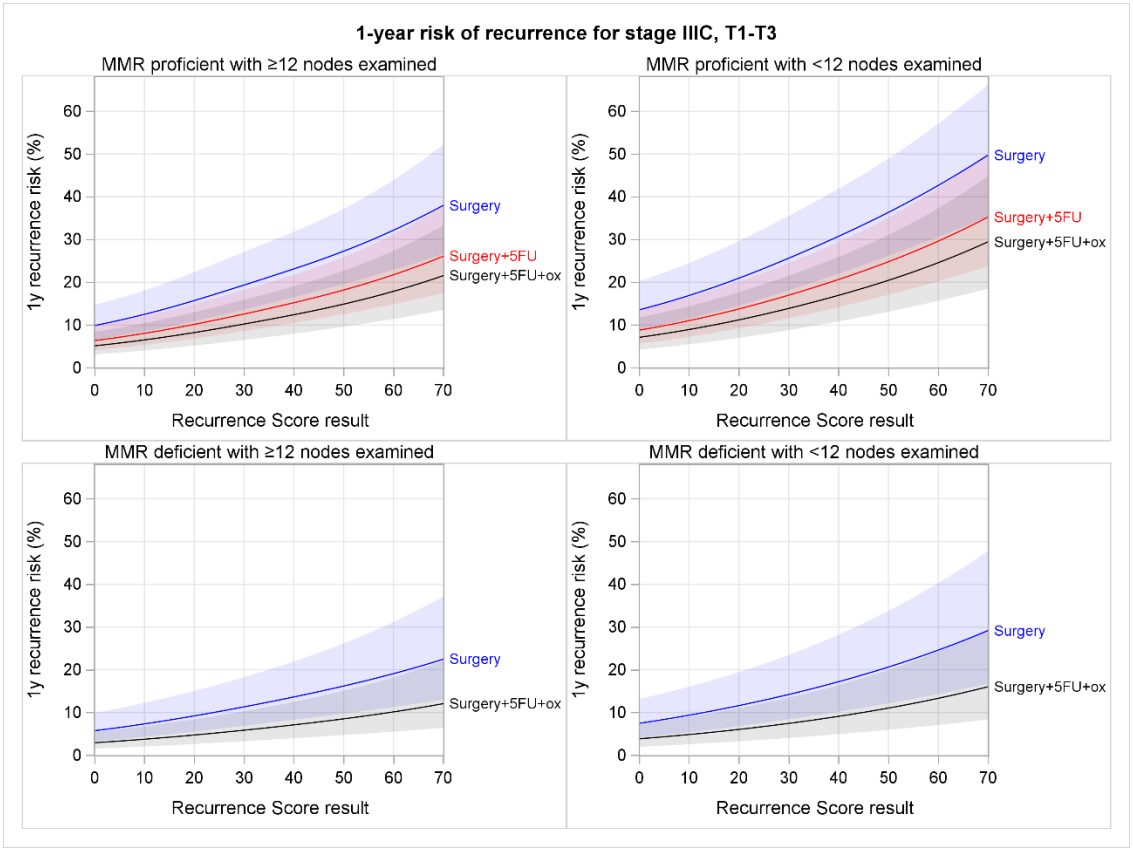
# 1-year Recurrence Risk Estimates

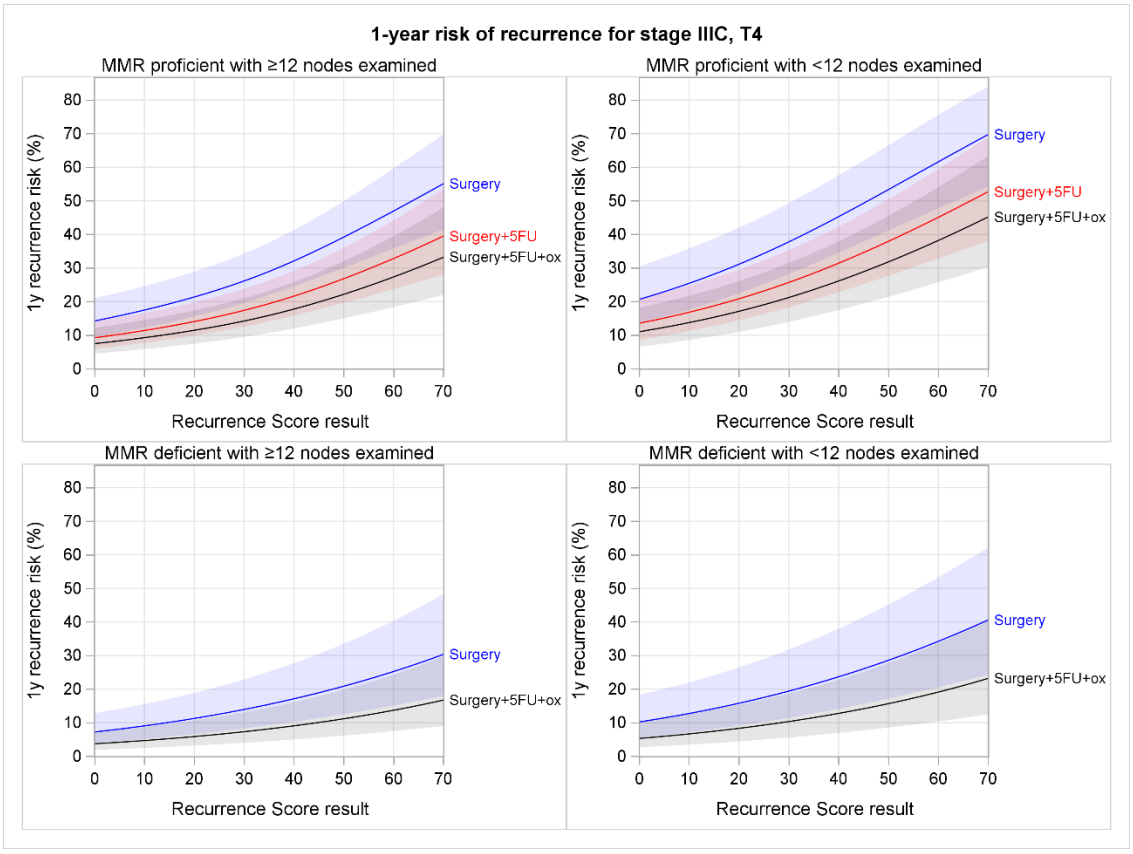

1-year risk of recurrence for stage II, T3
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
1y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox


1-year risk of recurrence for stage II, T4
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
1y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox


1-year risk of recurrence for stage IIIAB, T1-T3
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
1y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox


1-year risk of recurrence for stage IIIAB, T4
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
1y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox


1-year risk of recurrence for stage IIIC, T1-T3
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
1y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox


1-year risk of recurrence for stage IIIC, T4
MMR proficient with ≥12 nodes examined
MMR proficient with <12 nodes examined
MMR deficient with ≥12 nodes examined
MMR deficient with <12 nodes examined
1y recurrence risk (%)
Recurrence Score result
Surgery
Surgery+5FU
Surgery+5FU+ox

# APPENDIX. PRE-PLANNED STATISTICAL METHODS.

**Planned Statistical Methods for Patient-Specific Meta-Analysis of CALGB 9581, the Sunrise Study and NSABP C-07 for Prognosis for Recurrence in N-Stage II and III Colon Cancer Patients Receiving Surgery Alone, Surgery with 5FU and Surgery with 5FU and Oxaliplatin**

30 December 2020

***Studies***

The meta-analysis will use three validation studies for the Oncotype DX Colon Cancer Recurrence Score® (RS):

1. The parent CALGB 9581 study randomly assigned 1,713 patients with N-stage II colon cancer to treatment with edrecolomab or observation and found no survival difference. Venook *et al.* (2013) reported a prospective-retrospective study of the association of the Recurrence Score result with recurrence using a stratified cohort sample consisting of all patients with available tissue and recurrence (n=162) and a random (approximately 1:3) selection of patients without recurrence for a total sample size of 690 patients. The primary endpoint was recurrence in the first 5 years, with patients who died due to other causes without recurrence censored at last follow-up. CALGB 9581 enrolled patient from 1997 through 2002.
2. The Sunrise Study (Yamanaka *et al.* 2016) used a stratified cohort sample from 1,487 consecutive patients from 2000 through 2005 with N-stage II or III disease who had surgery alone, with 630 patients sampled for inclusion with a 1:2 ratio of recurrence and nonrecurrence. Sampling was stratified by N-stage (II vs. III). A total of 597 of the 630 patients were evaluable for analysis, 202 of whom experienced recurrences. The primary endpoint was time from surgery to first recurrence of colon cancer or death with a documented recurrence at the time of death. Patients who died before recurrence was observed were considered censored at last follow-up.
3. The National Surgical Adjuvant Breast and Bowel Project (NSABP) C-07 study randomly assigned patients with N-stage II and III colon cancer to fluorouracil (5FU) or 5FU plus oxaliplatin. Yothers *et al.* (2013) reported a prospective-retrospective study of the association of the Recurrence Score result with recurrence in 892 randomly selected patients from this study (50% of patients with available tissue), 245 of whom experienced recurrence. Among these patients, 449 had been randomly assigned to 5FU and 432 to

5FU with oxaliplatin. The primary endpoint was time to recurrence. Patients who died without recurrence were considered censored at the time of death. NSABP C-07 enrolled patients from 2000 through 2002.

The studies will be designated as follows:

Study 1: CALGB 9581

Study 2: Sunrise

Study 3: NSABP C-07

***Model Fitting***

The Cox proportional hazards regression models for each study will include the following terms:

1. The RS result as a continuous measure, fit as a linear term.
2. Number of nodes examined (<12 vs. ≥12).
3. T-stage T4 vs. T3 or less.
4. MMR status (deficient vs. proficient/unknown).
5. (NSABP C-07 and Sunrise only) N-stage (II, IIIA/B or IIIC).
6. (NSABP C-07 only) Oxaliplatin + 5FU vs. placebo + 5FU

Studies 1 (CALGB 9581) and 2 (Sunrise) used stratified cohort sampling (Gray 2009), so in the analysis of these studies each patient will be weighted in the analysis by the inverse sampling fraction in the patient's sampling stratum, and the covariance matrix of the regression parameter estimators variance will be estimated using the method of Lin and Wei (1989).

The analysis data set for each study will include all patients with non-missing values for the covariates RS, number of nodes examined, T-stage and N-stage. Patients with unknown MMR status will be combined with MMR proficient patients for analysis.

***Recurrence Risk Estimation***

The analysis requires an estimate of the hazard ratio for 5FU treatment added to surgery versus surgery alone. We will estimate this hazard ratio using a meta-analysis of the original QUASAR study (QUASAR study group, 2007) and a pooled analysis of NSABP trials (Wilkinson *et al.* 2010).

The log-rank observed-minus-expected ($O - E$) statistic, and its variance $V$, were reported for recurrence in the original trial (Quasar study group 2007). From these quantities, the log-rank statistic $Z = (O-E)/\sqrt{V}$ can be computed. Since patients were allocated to treatment with 5FU or observation with equal probability, the log hazard ratio can be estimated using the method of Schoenfeld (1981) by $Z\sqrt{4/D}$, where $D$ is the total number of recurrence events. The variance of this estimate is consistently estimated by $4/D$. The results of this calculation are in Table 1.

Table 1. Log-Rank Statistics and Estimate of Log Hazard Ratio from the QUASAR trial

| Events/Patients | | Log-rank Statistics | | Log Hazard Ratio (5FU:observation) | |
|---|---|---|---|---|---|
| 5FU | Observation | $O - E$ | $V$ | Estimate | Variance |
| 293 / 1622 | 359 / 1617 | −40.9 | 162.9 | −0.251 | 0.00613 |

Wilkinson *et al.* (2010) provide an estimate of the hazard ratio for recurrence for 5FU plus surgery versus surgery alone based on a pooled analysis of NSABP trials. The hazard ratio estimate from this analysis is 0.64 with 95% confidence interval (0.55, 0.74). The log hazard ratio estimate is thus $\ln 0.64 = -0.30111$ with an estimated standard error of $(\ln 0.74 - \ln 0.55)/\{2\Phi(0.975)\} = 0.075698.$

Combining the QUASAR and Wilkinson log hazard ratio estimates in a meta-analysis using inverse-variance weighting gives a 5FU log hazard ratio estimate of $\hat{\lambda}_{FU} = -0.35194$ with standard error $\hat{\sigma}_{\lambda_{5FU}} = 0.05444.$ and variance $\hat{\sigma}^2_{\lambda_{5FU}} = 0.002962.$ This corresponds to a hazard ratio for surgery and 5FU versus surgery alone of 0.703 with 95% confidence interval (0.632, 0.782).

The risks of recurrence at 1, 3 and 5 years after surgery will be estimated using patient-specific meta-analysis with special populations (Crager and Tang 2014), integrated with the meta-analysis 5FU treatment effect log hazard ratio. Here the special populations (not common to all

studies) are N-stage IIIA/B and IIIC patients and patients treated with oxaliplatin. The risk estimates will be constructed as follows.

Define the vector of covariates as

$$\mathbf{z} = \left( RS, I_{\{<12 \text{ nodes ex.}\}}, I_{\{T4\}}, I_{\{MMRD\}}, I_{\{IIIA/B\}}, I_{\{IIIC\}}, I_{\{\text{oxali}\}} \right)^{\mathrm{T}}$$

and define

$$\mathbf{z}_0 = \left( RS, I_{\{<12 \text{ nodes ex.}\}}, I_{\{T4\}}, I_{\{MMRD\}}, 0, 0, 0 \right)^{\mathrm{T}},$$
$$\mathbf{z}_{IIIA/B} = \left( 0, 0, 0, 0, 1, 0, 0 \right)^{\mathrm{T}},$$
$$\mathbf{z}_{IIIC} = \left( 0, 0, 0, 0, 0, 1, 0 \right)^{\mathrm{T}}$$

and

$$\mathbf{z}_{\text{oxali}} = \left( 0, 0, 0, 0, 0, 0, 1 \right)^{\mathrm{T}}.$$

Since the overall recurrence risk has decreased over time, we will use the events from the latest-enrolling two studies (studies 2 and 3) to estimate the baseline cumulative hazard, with risk modification for individual presenting patients based on the regression parameters from each study. Since study 1 enrolled only N-stage II patients, the baseline for this study will be estimated using only the N-stage II patients in studies 2 and 3. Similarly, since no patient in studies 1 or 2 was treated with oxaliplatin, the baselines for those studies estimated using study 3 will be based on patients not treated with oxaliplatin. For each study $k = 1, 2, 3,$ define $I_i^{(k,2)}$ as the indicator for whether study 2 patient $i$ is included in the baseline cumulative hazard estimator using study 2. Define the indicators $I_i^{(k,3)}$ similarly for baseline cumulative hazard estimators using study 3. Let $\hat{\boldsymbol{\beta}}_k = \left( \beta_1^{(k)}, \hat{\beta}_2^{(k)}, \ldots, \hat{\beta}_7^{(k)} \right)^{\mathrm{T}}$ and $\hat{\mathbf{V}}_k$ be the estimated proportional hazards regression parameter vector and its estimated covariance matrix for study $k = 1, 2, 3.$ Set $\hat{\beta}_5^{(1)} = 0, \hat{\beta}_6^{(1)} = 0$ and $\hat{\beta}_7^{(1)} = 0$ and set $\hat{\beta}_7^{(2)} = 0,$ and set all corresponding elements of $\hat{\mathbf{V}}_k$ to 0. Let $\mathbf{z}_i^{(k)}$ be the observed covariate vector for patient $i = 1, 2, \ldots, n_k$ in study *k*. The Breslow-method estimator of the baseline cumulative hazard function at time *T* using the regression coefficients for study *k* and the events for study 2 is

$$\hat{\Lambda}_0^{(k,2)}(T) = \int_0^T \frac{I_i^{(k,2)} s_i^{(2)} \, \mathrm{d}\, N^{(2)}(t)}{\sum_{i=1}^{n_2} I_i^{(k,2)} s_i^{(2)} Y_i^{(2)}(t) \exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}} \mathbf{z}_i^{(2)}\right)},$$

where $N^{(2)}(t)$ is the event-counting process for study 2, $s_i^{(2)}$ is the stratified cohort sampling weight, and $Y_i^{(2)}(t)$ is the indicator for whether patient $i$ in study 2 is in the risk set at time *t*.

The variance of $\hat{\Lambda}_0^{(k,2)}(T)$ due to the event count is consistently estimated by

$$\widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,2)}(T)\right\} = \int_0^T \frac{I_i^{(k,2)} \left(s_i^{(2)}\right)^2 \mathrm{d}\, N^{(2)}(t)}{\left\{\sum_{i=1}^{n_3} I_i^{(k,2)} s_i^{(2)} Y_i^{(2)}(t) \exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}} \mathbf{z}_i^{(2)}\right)\right\}^2}.$$

Since all patients in study 3 received 5FU in addition to surgery, the baseline cumulative hazard estimator for study *k* using the events from study 3 is

$$\hat{\Lambda}_0^{(k,3)}(T) = \int_0^T \frac{I_i^{(k,3)} \, \mathrm{d}\, N^{(3)}(t)}{\sum_{i=1}^{n_3} I_i^{(k,3)} Y_i^{(3)}(t) \exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}} \mathbf{z}_i^{(3)} + \hat{\lambda}_{5FU}\right)}.$$

The variance of $\hat{\Lambda}_0^{(k,3)}(T)$ due to the event count is consistently estimated by

$$\widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,3)}(T)\right\} = \int_0^T \frac{I_i^{(k,3)} \, \mathrm{d}\, N^{(3)}(t)}{\left\{\sum_{i=1}^{n_3} I_i^{(k,3)} Y_i^{(3)}(t) \exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}} \mathbf{z}_i^{(3)} + \hat{\lambda}_{5FU}\right)\right\}^2}.$$

A fixed effects meta-analysis baseline cumulative hazard estimator combining studies 2 and 3 is

$$\hat{\Lambda}_0^{(k,2+3)}(T) = \omega_2^{(\Lambda_0)} \hat{\Lambda}_0^{(k,2)}(T) + \omega_3^{(\Lambda_0)} \hat{\Lambda}_0^{(k,3)}(T),$$

where

$$\omega_{k,2}^{(\Lambda_0)} = \frac{1 \Big/ \widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,2)}(T)\right\}}{1 \Big/ \widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,2)}(T)\right\} + 1 \Big/ \widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,3)}(T)\right\}}$$

and

$$\omega_{k,3}^{(\Lambda_0)} = \frac{1 \Big/ \widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,3}(T)\right\}}{1 \Big/ \widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,2)}(T)\right\} + 1 \Big/ \widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,3)}(T)\right\}}.$$

Let $I_{5FU}$ and $I_{\mathrm{oxali}}$ be the indicators for whether the patient is to receive 5FU and oxaliplatin (with 5FU). For an individual patient with covariate vector $\mathbf{z}$, the estimated natural logarithm of the cumulative hazard at time *T* for study $k = 1, 2, 3$ is

$$\hat{\rho}_k(T;\mathbf{z}) = \hat{\boldsymbol{\beta}}_k^{\mathrm{T}}\mathbf{z} + I_{5FU}\hat{\lambda}_{5FU} + \ln\hat{\Lambda}_0^{(k,2+3)}(T).$$

Defining the gradient operator $\nabla_{\hat{\boldsymbol{\beta}}_k} = \left(\partial/\partial\beta_1^{(k)}, \partial/\partial\beta_2^{(k)}, ,\ldots, \partial/\partial\beta_7^{(k)}\right)^{\mathrm{T}}$, setting the elements for regression parameters that do not exist in each study to 0, we have

$$\nabla_{\hat{\boldsymbol{\beta}}_k}\hat{\rho}_k(T;\mathbf{z}) = \mathbf{z} + \frac{\nabla_{\hat{\boldsymbol{\beta}}_k}\hat{\Lambda}_0^{(k,2+3)}(T)}{\hat{\Lambda}_0^{(k,2+3)}(T)},$$

with $\nabla_{\hat{\boldsymbol{\beta}}_k}\hat{\Lambda}_0^{(k,2+3)}(T) = -\boldsymbol{\gamma}^{(k,2+3)}(T)$, where

$$\boldsymbol{\gamma}^{(k,2+3)}(T) = \omega_{k,2}^{(\Lambda_0)}\int_0^T \frac{\sum_{i=1}^{n_2} I_i^{(k,2)} s_i^{(2)} Y_i^{(2)}(t)\exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}}\mathbf{z}_i^{(2)}\right)\mathbf{z}_i^{(2)}}{\sum_{i=1}^{n_2} I_i^{(k,2)} s_i^{(2)} Y_i^{(3)}(t)\exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}}\mathbf{z}_i^{(2)}\right)}\mathrm{d}\hat{\Lambda}_0^{(k,2)}(t)$$
$$+ \omega_{k,3}^{(\Lambda_0)}\int_0^T \frac{\sum_{i=1}^{n_3} I_i^{(k,3)} Y_i^{(3)}(t)\exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}}\mathbf{z}_i^{(3)} + \hat{\lambda}_{5FU}\right)\mathbf{z}_i^{(3)}}{\sum_{i=1}^{n_3} I_i^{(k,3)} Y_i^{(3)}(t)\exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}}\mathbf{z}_i^{(3)} + \hat{\lambda}_{5FU}\right)}\mathrm{d}\hat{\Lambda}_0^{(k,3)}(t).$$

Also, we have

$$\frac{\partial\hat{\rho}_k(T;\mathbf{z})}{\partial\hat{\lambda}_{5FU}} = I_{5FU} + \frac{\partial\hat{\Lambda}_0^{(k,2+3)}(T)}{\partial\hat{\lambda}_{5FU}}\Big/\hat{\Lambda}_0^{(k,2+3)}(T),$$

with

$$\frac{\partial\hat{\Lambda}_0^{(k,2+3)}(T)}{\partial\hat{\lambda}_{5FU}} = -\omega_{k,3}^{(\Lambda_0)}\int_0^T \frac{\sum_{i=1}^{n_3} Y_i^{(3)}(t)\exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}}\mathbf{z}_i^{(3)} + \hat{\lambda}_{5FU}\right)}{\sum_{i=1}^{n_3} Y_i^{(3)}(t)\exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}}\mathbf{z}_i^{(3)} + \hat{\lambda}_{5FU}\right)}\mathrm{d}\hat{\Lambda}_0^{(k,3)}(t)$$
$$= -\omega_{k,3}^{(\Lambda_0)}\hat{\Lambda}_0^{(k,3)}(T)$$

so that

$$\frac{\partial\hat{\rho}_k(T;\mathbf{z})}{\partial\hat{\lambda}_{5FU}} = I_{5FU} - \omega_{k,3}^{(\Lambda_0)}\frac{\hat{\Lambda}_0^{(k,3)}(T)}{\hat{\Lambda}_0^{(k,2+3)}(T)}.$$

Therefore, for $k = 1, 2, 3$, the variance of $\hat{\rho}_k(T;\mathbf{z})$ is consistently estimated by

$$\hat{\sigma}_k^2(T;\mathbf{z}) = \left(\nabla_{\hat{\boldsymbol{\beta}}_k}\hat{\rho}_k(T;\mathbf{z})\right)^{\mathrm{T}}\hat{\mathbf{V}}_k\left(\nabla_{\hat{\boldsymbol{\beta}}_k}\hat{\rho}_k(T;\mathbf{z})\right) + \left\{\frac{\partial\hat{\rho}_k(T;\mathbf{z})}{\partial\hat{\lambda}_{5FU}}\right\}^2\hat{\sigma}_{\lambda_{5FU}}^2$$
$$+ \frac{\left\{\omega_{k,2}^{(\Lambda_0)}\right\}^2\widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,2)}(T)\right\} + \left\{\omega_{k,3}^{(\Lambda_0)}\right\}^2\widehat{\mathrm{Var}}_N\left\{\hat{\Lambda}_0^{(k,3)}(T)\right\}}{\left\{\hat{\Lambda}_0^{(k,2+3)}(T)\right\}^2}.$$

Using the special population PSMA method (Crager and Tang 2014) to adjust for the effects of N-stage and oxaliplatin treatment, a fixed effects PSMA log cumulative hazard estimate for recurrence during the first $T$ = 1, 3 or 5 years is

$$\hat{\rho}\left(T;\mathbf{z}\right)=\sum_{k=1}^{3}\omega_k\left(\mathbf{z}_0\right)\hat{\rho}_k\left(T;\mathbf{z}^{\dagger}\right),$$

where

$$\mathbf{z}^{\dagger}=\mathbf{z}_0+\frac{I_{\{IIIA/B\}}\mathbf{z}_{IIIA/B}+I_{\{IIIC\}}\mathbf{z}_{IIIC}}{\omega_2(\mathbf{z}_0)+\omega_3(\mathbf{z}_0)}+\frac{I_{\text{oxali}}\mathbf{z}_{\text{oxali}}}{\omega_3(\mathbf{z}_0)}$$

and

$$\omega_k(\mathbf{z}_0)=\frac{w_k(5;\mathbf{z}_0)}{\sum_{j=1}^{3}w_j(5;\mathbf{z}_0)}$$

with

$$w_k\left(5;\mathbf{z}_0\right)=\frac{1}{\hat{\sigma}_k^2(5;\mathbf{z}_0)}.$$

The time is set at 5 years for the calculation of the weights so that the 1-, 3- and 5-year risk estimates will be mutually consistent.

The gradient of $\hat{\rho}\left(T;\mathbf{z}\right)$ with respect to $\hat{\boldsymbol{\beta}}_k^{\mathrm{T}}, k=1,2,3,$ is

$$\nabla_{\hat{\boldsymbol{\beta}}_k}\hat{\rho}\left(T;\mathbf{z}\right)=\omega_k(\mathbf{z}_0)\nabla_{\hat{\boldsymbol{\beta}}_k}\hat{\rho}_k\left(T;\mathbf{z}^{\dagger}\right).$$

Letting $t_i^{(2)}$ denote the time to event or censoring and $N_i^{(2)}\left(t\right)$ denote the event-counting process for patient $i=1,2,\ldots,n_2$ in study 2, the partial derivative of $\hat{\Lambda}_0^{(k,2)}(T)$ with respect to $\mathrm{d}\,N_i^{(2)}\left(t_i^{(2)}\right)$ is

$$\frac{\partial\hat{\Lambda}_0^{(k,2)}(T)}{\partial\,\mathrm{d}\,N_i^{(2)}\left(t_i^{(2)}\right)}=I_{\left\{t_i^{(2)}\le T\right\}}\frac{I_i^{(k,2)}s_i^{(2)}\,\mathrm{d}\,N_i^{(2)}\left(t_i^{(2)}\right)}{\sum_{j=1}^{n_2}I_i^{(k,2)}s_i^{(2)}Y_j^{(2)}\left(t_i^{(2)}\right)\exp\left(\hat{\boldsymbol{\beta}}_k^{\mathrm{T}}\mathbf{z}_j^{(2)}\right)}=I_{\left\{t_i^{(2)}\le T\right\}}\,\mathrm{d}\,\hat{\Lambda}_0^{(k,2)}\left(t_i^{(2)}\right),$$

where $I_{\left\{t_i^{(2)}\le T\right\}}$ is the indicator function for $t_i^{(2)}\le T,$ and $\mathrm{d}\,\hat{\Lambda}_0^{(k,2)}\left(t_i^{(2)}\right)$ is the increment in the baseline cumulative hazard estimate at time $t_i^{(2)}$. Therefore,

$$\frac{\partial\hat{\rho}(T;\mathbf{z})}{\partial\,\mathrm{d}\,N_i^{(2)}\left(t_i^{(2)}\right)}=\sum_{k=1}^{3}\omega_{k,2}^{(\Lambda_0)}\omega_k\left(\mathbf{z}_0\right)I_{\left\{t_i^{(2)}\le T\right\}}\frac{\mathrm{d}\,\hat{\Lambda}_0^{(k,2)}\left(t_i^{(2)}\right)}{\hat{\Lambda}_0^{(k,2)}(T)}.$$

Similarly, letting $t_i^{(3)}$ denote the time to event or censoring and $N_i^{(3)}(t)$ denote the event-counting process for patient $i = 1, 2, \ldots, n_3$ in study 3,

$$\frac{\partial \hat{\rho}(T;\mathbf{z})}{\partial \, \mathrm{d} N_i^{(3)}\left(t_i^{(3)}\right)} = \sum_{k=1}^{3} \omega_{k,3}^{(\Lambda_0)} \omega_k\left(\mathbf{z}_0\right) I_{\left\{t_i^{(3)} \le T\right\}} \frac{\mathrm{d}\, \hat{\Lambda}_0^{(k,3)}\left(t_i^{(3)}\right)}{\hat{\Lambda}_0^{(k,3)}(T)}.$$

Since the number and timing of events are asymptotically independent of the regression parameter estimates (Tsiatis 1981), assuming the three studies represent independent samples of patients and using methods similar to those in Therneau and Grambsch (2000, Section 2.1), the variance of $\hat{\rho}(T;\mathbf{z})$ is consistently estimated by

$$\begin{aligned}\widehat{\mathrm{Var}}\left\{\hat{\rho}(T;\mathbf{z})\right\} = &\sum_{k=1}^{3}\left(\omega_k(\mathbf{z}_0)\nabla_{\hat{\boldsymbol{\beta}}_k}\hat{\rho}_k\left(T;\mathbf{z}^{\dagger}\right)\right)^{\mathrm{T}} \hat{\mathbf{V}}_k \left(\omega_k(\mathbf{z}_0)\nabla_{\hat{\boldsymbol{\beta}}_k}\hat{\rho}_k\left(T;\mathbf{z}^{\dagger}\right)\right) \\ &+ \left\{\sum_{k=1}^{3}\omega_k(\mathbf{z}_0)\frac{\partial\hat{\rho}_k\left(\mathbf{z}^{\dagger}\right)}{\partial\hat{\lambda}_{5FU}}\right\}^2 \hat{\sigma}^2_{\lambda_{5FU}} \\ &+ \sum_{i=1}^{n_2}\left\{\frac{\partial\hat{\rho}(T;\mathbf{z})}{\partial\,\mathrm{d}N_i^{(2)}\left(t_i^{(2)}\right)}\right\}^2 \mathrm{d}N_i^{(2)}\left(t_i^{(2)}\right) \\ &+ \sum_{i=1}^{n_3}\left\{\frac{\partial\hat{\rho}(T;\mathbf{z})}{\partial\,\mathrm{d}N_i^{(3)}\left(t_i^{(3)}\right)}\right\}^2 \mathrm{d}N_i^{(3)}\left(t_i^{(3)}\right).\end{aligned}$$

The third term in the sum above can be written

$$\sum_{i=1}^{n_2}\left\{\frac{\partial\hat{\rho}\left(T;\mathbf{z}\right)}{\partial\,\mathrm{d}N_i^{(2)}\left(t_i^{(2)}\right)}\right\}^2 \mathrm{d}N_i^{(2)}\left(t_i^{(2)}\right) = \left\{\mathbf{c}_2(\mathbf{z}_0)\right\}^{\mathrm{T}} \mathbf{X}_2(T)\left\{\mathbf{c}_2(\mathbf{z}_0)\right\},$$

where $\mathbf{c}_2(\mathbf{z}_0) = \left(\omega_{1,2}^{(\Lambda_0)}\omega_1\left(\mathbf{z}_0\right), \omega_{2,2}^{(\Lambda_0)}\omega_2\left(\mathbf{z}_0\right), \omega_{3,2}^{(\Lambda_0)}\omega_3\left(\mathbf{z}_0\right)\right)^{\mathrm{T}}$ and $\mathbf{X}_2(T)$ is the matrix with element $x_{kl}^{(2)}(T)$ in row *k* and column *l*, with

$$x_{kl}^{(2)}(T) = \int_0^T \frac{\mathrm{d}\,\hat{\Lambda}_0^{(k,2)}\left(t\right)}{\hat{\Lambda}_0^{(k,2)}\left(T\right)} \frac{\mathrm{d}\,\hat{\Lambda}_0^{(l,2)}\left(t\right)}{\hat{\Lambda}_0^{(l,2)}\left(T\right)} \mathrm{d}N^{(2)}\left(t\right).$$

Similarly, the fourth term in (2) can be written

$$\sum_{i=1}^{n_3}\left\{\frac{\partial\hat{\rho}\left(T;\mathbf{z}\right)}{\partial\,\mathrm{d}N_i^{(3)}\left(t_i^{(3)}\right)}\right\}^2 \mathrm{d}N_i^{(3)}\left(t_i^{(3)}\right) = \left\{\mathbf{c}_3(\mathbf{z}_0)\right\}^{\mathrm{T}} \mathbf{X}_3(T)\left\{\mathbf{c}_3(\mathbf{z}_0)\right\},$$

where $\mathbf{c}_3(\mathbf{z}_0) = \left(\omega_{1,3}^{(\Lambda_0)}\omega_1(\mathbf{z}_0), \omega_{2,3}^{(\Lambda_0)}\omega_2(\mathbf{z}_0), \omega_{3,3}^{(\Lambda_0)}\omega_3(\mathbf{z}_0)\right)^{\mathrm{T}}$ and $\mathbf{X}_3(T)$ is matrix with element $x_{kl}^{(3)}(T)$ in row *k* and column *l*, with

$$x_{kl}^{(3)}(T) = \int_0^T \frac{\mathrm{d}\hat{\Lambda}_0^{(k,3)}(t)}{\hat{\Lambda}_0^{(k,3)}(T)} \frac{\mathrm{d}\hat{\Lambda}_0^{(l,3)}(t)}{\hat{\Lambda}_0^{(l,3)}(T)} \mathrm{d}N^{(3)}(t).$$

Transforming to the risk scale, the estimated risk of a recurrence by time *T* is $\hat{r}(T;\mathbf{z}) = 1 - \exp\left[-\exp\{\hat{\rho}(T;\mathbf{z})\}\right]$. A level-$\alpha$ confidence interval for the recurrence risk in the first *T* years after surgery has endpoints $1 - \exp\left(-\exp\left[\hat{\rho}(T;\mathbf{z}) \pm \Phi^{-1}(1-\alpha/2)\sqrt{\widehat{\mathrm{Var}}\{\hat{\rho}(T;\mathbf{z})\}}\right]\right)$, where $\Phi$ denotes the cumulative distribution function of the standard normal distribution.

***Testing Analysis Assumptions***

The assumption that there is no interaction among the 5 prognostic factors (RS result, N-stage, T-stage, number of nodes examined and MMR proficiency) will be tested using meta-analysis Wald tests. For each of the 4 categorical factors, a model will be fit to each study allowing different regression parameters across the levels of the categorical factor. A fixed-effect meta-analysis estimate for the difference(s) across levels in these parameters will be constructed using inverse variance weighting and the statistical significance of the resulting meta-analysis difference estimates will be assessed using a Wald test at a significance level of 0.10. The potential for interaction of the number of nodes examined with stage and MMR proficiency with stage will be of particular interest. If the presence of an interaction is indicated, a model including that interaction will be considered for generating the recurrence risk estimates.

***References***

Crager MR, Tang G (2014). Patient-specific meta-analysis for risk assessment using multivariate proportional hazards regression. *Journal of Applied Statistics* **41**:2676–2695.

Gray R (2009). Weighted analysis for cohort sampling designs. *Lifetime Data Analysis*, **15**:24–40.

Lin DY, Wei LJ (1989). The robust inference for the Cox proportional hazards model. *Journal of the American Statistical Association* **84**:1074–1078.

QUASAR Study Group (2007). Adjuvant chemotherapy versus observation in patients with colorectal cancer: a randomised study. *Lancet* **370**:2020–2029.

Schoenfeld D (1981). The asymptotic properties of nonparametric tests for comparing survival distributions. *Biometrika* **68**:316–319.

Therneau TM, Grambsch PM (2000). *Modeling Survival Data: Extending the Cox Model*. New York: Springer.

Tsiatis A (1981). A large sample study of the estimates for the integrated hazard function in Cox's regression model for survival data. *Annals of Statistics* **9**:93–108.

Venook AP, Neidzwiecki D, Lopatin M, Ye X, Lee M, Friedman PN, Frankel W, Clark-Langone K, Millward C, Shak S, Goldberg RM, Mahmoud NN, Waren RS, Schilsky RL, Bertagnolli MM (2013). Biologic determinants of tumor recurrence in Stage II colon cancer: validation study of the 12-gene recurrence score in cancer and Leukemia Group B (CALGB) 9581. *Journal of Clinical Oncology* **31**:1775–1781.

Wilkinson NW, Yothers G, Lopa S, Costantino JP, Petrelli NJ, Wolmark N (2010). Long-term survival results of surgery alone versus surgery plus 5-fluorouracil and leucovorin for Stage II and Stage III colon cancer: pooled analysis of NSABP C-01 through C-05. A baseline from which to compare modern adjuvant trials. *Annals of Surgical Oncology* 17:959–966.

Yamanaka T, Oki E, Yamazaki K, Yamaguchi K, Muro K, Uetake H, Sato T, Nishina T, Ikeda M, Kato T, Kanazawa A, Kusumoto T, Chao C, Lopatin M, Krishnakumar J, Bailey H, Akagi K, Ochiai A, Ohtsu A, Ohashi Y, Yoshino T (2016). 12-gene recurrence score assay stratifies the recurrence risk in Stage II/III colon cancer with surgery alone: the SUNRISE study. *Journal of Clinical Oncology* **34**:2906–2913.

Yothers G, O'Connell MJ, Lee M, Lopatin M, Clark-Langone KM, Millward C, Paik S, Sharif S, Shak S, Wolmark N (2013). Validation of the 12-gene colon cancer recurrence score in NSABP C-07 as a predictor of recurrence in patients with Stage II and III colon cancer treated with fluorouracil and leucovorin (FU/LV) and FU/LV plus oxaliplatin. *Journal of Clinical Oncology* **31**:4512–4519.