Updating Colorectal Cancer (CRC) Survival in the CRC-SPIN Micro-simulation Model

Authors: Tracey Marsh MSc, Chester Pabiniak MSc, Eric Johnson MSc
Investigator: Carolyn Rutter, PhD
Affiliations: Group Health Research Institute

Additive Hazards Model

Annual hazard rates:
All cause mortality is modeled as the sum of the force of mortality due to CRC and the force of mortality due to all other causes:

$$\lambda_i(s) = \lambda_i^{CRC}(s) + \lambda_i^{-CRC}(s)$$

Interval specific survival rates:

$$p_i^D = \exp \left[ - \int_{i-1}^{i} \lambda_i^{D}(s)ds \right]$$

D indicates: All causes, CRC, or non-CRC causes.

Implementation within CRC-SPIN

Circles represent states (C# classes) that describe (implement) Natural History Model Processes.

Relative Survival

GLM with complementary log-log link function:

$$\ln \left[ \frac{-\ln \left( \frac{p_i}{p_i^{CRC}} \right)}{1 - p_i^{CRC}} \right] = \ln \left[ \int_{i-1}^{i} \lambda^{CRC}(s)ds \right]$$

Linear on the log cumulative hazard (interval increment) scale.

Binomial error function:

$$N_i^{CRC} \left[ \frac{1 - p_i^{CRC}}{1 - p_i} \right]$$

Estimating: number of deaths among individuals diagnosed with CRC relative to the number of expected deaths in a matched population in each one-year follow-up interval.

Executed in SAS using a call to PROC GENMOD


Secular Trends

For any number of years since diagnosis, i:

$$\ln \left( \Lambda_i^{CRC} \right) = \Lambda_i^{CRC} - Z_i \gamma + X\beta$$

Complete hazard increments are estimated:

X baseline covariates
i index of years since CRC diagnosis
Z time-varying covariates

Model stratified by:
Location (colon, rectum) and Stage of CRC
Covariates included in model:
Age group at diagnosis
Sex
Year of CRC diagnosis
Interaction between Age & Years since Dx

Natural History Model

Simulates all modeled aspects of an individual, subject to individual and population-level constraints in the absence of any screening program. Screening is modeled separately and modifies simulated natural histories.

Years until non-CRC death is simulated at the beginning of the natural history process.

Years until CRC death is determined at time of diagnosis which can be either at screening (pre-clinical cancerous lesions) or by presenting clinically (lesions in clinical disease state).

Example of inverse CDF lookup algorithm:

1) Results from the relative survival analysis are saved in formatted XML files. These files are read by the CRC-SPIN C# model and used to generate CRC survival curves.

2) Time from diagnosis to colorectal cancer death is calculated through an inverse CDF lookup algorithm:

   • An individual’s percentile within the population survival is directly simulated (sample from a uniform distribution).
   • Years until death is the first year that the probability of survival falls below the percentile.
   • The additional fraction of the year survived is another draw from a standard uniform distribution.

3) Simulated date of death = min(CRC dod, non-CRC dod)

Data Sources

• All (other) cause survival
  National Center for Health Statistics

• Observed survival in population with CRC
  SEER (1975-2003)
  individuals with CRC Dx
  with follow-up through 2009
  black or white race

• Expected survival in non-CRC population
  SEER*Stat (NCHS)
  only available for black and white races