

September 12, 2008

Matching Results for Invasive Colorectal Cancer in the WHI Observational Study

Case / Control selection:

All centrally or locally (if no central adjudication) confirmed cases of invasive colorectal cancer were selected as potential cases from the database as of April 24, 2008. Control participants were required to be free of any type of colorectal cancer, invasive or non-invasive. Participants with centrally denied outcomes were allowed into the control pool.

Cases were only excluded if they had no available baseline EDTA, baseline serum, year 3 serum, baseline red blood cells, and DNA. Essentially, if a case had at least one of these components they remained in the matching set. Case participants with a genetic refusal were still matched for bloods. Controls were excluded if they had inadequate EDTA, baseline serum, red blood cells, or DNA. Controls were also excluded if they had refusal for genetic studies. When attempting matching to those cases with year 3 serum available, those controls without year 3 serum were removed from the control pool.

In an effort to merge with earlier work, a tiered structure was used in matching to potential control sets. First priority was to match to existing hip controls from BAA03-Genome wide association study to identify genetic components of hip fracture, PI, Rebecca Jackson. Once this was done the next matching was done to use as many controls as possible from the AS206-Effects of E-alone HT on body composition (PI-Riki Peters). Once these two sources of controls were used, any remaining cases would be matched to controls from the general WHI population. At each control set level, matching was first done for those participants who had blood available at both baseline and year 3, matching as many as possible to controls who also had the year 3 serum. Matching was then done on cases who were missing either the baseline or year 3 blood components, only matching on the available draws. Finally, for the last few cases, the matching criteria were relaxed to ensure that every case would have a control. The combination of the different control sets and match types resulted in an 8-phase matching set up, with the phases being defined (in sequential order) as:

- 1) BAA 3 hip fracture controls matching both baseline and year 3 blood
- 2) BAA 3 hip fracture controls matching baseline blood only
- 3) BAA 3 hip fracture controls matching year 3 blood only
- 4) AS206 controls, matching both baseline and year 3 blood
- 5) AS206 controls, matching baseline blood only
- 6) WHI general population, matching both baseline and year 3 blood
- 7) WHI general population, matching baseline blood only
- 8) WHI general population, relaxed matching parameters

Matching criteria:

Matching was done on age at screening, enrollment date, baseline blood draw date (phases 1,2,4-8), year 3 blood draw date (phases 1,3,4,6,8), ethnicity (white, black, hispanic, other), hysterectomy status, and prevalent condition at baseline. Ethnicity, hysterectomy status, and prevalent condition status (the categorical variables) were matched exactly, and the continuous matching variables were selected based on a criteria to minimize an overall distance measure, with a distance of one year of

age equivalent to a distance of 45 days in either enrollment date or either draw date. Matching was done in a time-forward manner, selecting one control for each case from the risk set at the time of the case's event, thus ensuring that each control had at least as much control time as its corresponding case (Bergstralh EJ, Kosanke JL. Computerized matching of cases to controls. Technical Report #56, Department of Health Sciences Research, Mayo Clinic, Rochester MN. April 1995). SAS code is available to implement this matching scheme.

Matching summary:

Phase 1: BAA 3 hip fracture controls matching both baseline and year 3 blood

A total of 449 case / control pairs were successfully matched.

Matching Factor	Sum (weighted) of Absolute Differences Mean (min,Q25, Q50, Q75,max)	Cases Mean / N, %	Controls Mean / N, %
Age (years)	-0.33 (-3, -1, 0, 1, 3)	68.69	69.02
Enrollment date (days)	5.45 (-317, -26, 3, 34, 233)	13441.4	13435.9
Baseline blood draw date	0.16 (-180, -37, -2, 41, 177)	13397.7	13397.5
Year 3 blood draw date	3.32 (-179, -32, 2, 39, 180)	14565.7	14562.4
Ethnicity (1=white, 2=black, 3=hispanic, 4=other/unknown)	0	449 (100.0%) white	
Hysterectomy at baseline	0	163 (36.3%) yes	
History of colorectal cancer (0=no, 1=yes)	0	1 (0.2%) yes	

Phase 2: BAA 3 hip fracture controls matching baseline blood only

A total of 99 case / control pairs were successfully matched.

Matching Factor	Sum (weighted) of Absolute Differences Mean (min,Q25,Q75,max)	Cases Mean / N, %	Controls Mean / N, %
Age (years)	-0.10 (-3, -1, 0, 1, 3)	69.38	69.48
Enrollment date (days)	8.58 (-272, -38, -8, 51, 316)	13433.0	13424.4
Baseline blood draw date	4.52 (-170, -49, 0, 55, 182)	13385.2	13380.7
Ethnicity (1=white, 2=black, 3=hispanic, 4=other/unknown)	0	99 (100.0%) white	
Hysterectomy at baseline	0	39 (39.4%) yes	
History of colorectal cancer (0=no, 1=yes)	0	0 (0.0%) yes	

Phase 3: BAA 3 hip fracture controls matching year 3 blood only

A total of 1 case / control pair was successfully matched.

Matching Factor	Sum (weighted) of Absolute Differences Mean (min,Q25,Q75,max)	Cases Mean / N, %	Controls Mean / N, %
Age (years)	0.00 (0, 0, 0, 0, 0)	77.00	77.00
Enrollment date (days)	45.00 (45, 45, 45, 45, 45)	12852.0	12807.0
Year 3 blood draw date	28.00 (28, 28, 28, 28, 28)	13934.0	13906.0
Ethnicity (1=white, 2=black, 3=hispanic, 4=other/unknown)	0	1 (100.0%) white	
Hysterectomy at baseline	0	0 (0.0%) yes	
History of colorectal cancer (0=no, 1=yes)	0	0 (0.0%) yes	

Phase 4: AS206 controls, matching both baseline and year 3 blood

A total of 269 case / control pairs were successfully matched.

Matching Factor	Sum (weighted) of Absolute Differences Mean (min,Q25,Q75,max)	Cases Mean / N, %	Controls Mean / N, %
Age (years)	-0.14 (-3, -1, 0, 0, 3)	64.04	64.19
Enrollment date (days)	2.41 (-165, -14, 0, 14, 244)	13519.7	13517.3
Baseline blood draw date	0.43 (-171, -29, 0, 29, 165)	13478.4	13478.0
Year 3 blood draw date	12.30 (-178, -23, 8, 50, 175)	14651.4	14639.1
Ethnicity (1=white, 2=black, 3=hispanic, 4=other/unknown)	0	213 (79.2%) white 33 (12.3%) black 8 (3.0%) hispanic 15 (5.6%) other/unknown	
Hysterectomy at baseline	0	132 (49.1%) yes	
History of colorectal cancer (0=no, 1=yes)	0	8 (3.0%) yes	

Phase 5: AS206 Selenium controls, matching baseline blood only

A total of 76 case / control pairs were successfully matched.

Matching Factor	Sum (weighted) of Absolute Differences Mean (min,Q25,Q75,max)	Cases Mean / N, %	Controls Mean / N, %
Age (years)	-0.17 (-3, 0, 0, 0, 3)	65.12	65.29
Enrollment date (days)	1.82 (-194, -0.5, 11.5, 17.5, 71)	13526.1	13524.3
Baseline blood draw date	-8.89 (-178, -36, 5.5, 27, 158)	13476.3	13485.2
Ethnicity (1=white, 2=black, 3=hispanic, 4=other/unknown)	0	54 (71.1%) white 13 (17.1%) black 5 (6.6%) hispanic 4 (5.3%) other/unknown	
Hysterectomy at baseline	0	37 (48.7%) yes	
History of colorectal cancer (0=no, 1=yes)	0	1 (1.3%) yes	

Phase 6: WHI general population, matching both baseline and year 3 blood

A total of 69 case / control pairs were successfully matched.

Matching Factor	Sum (weighted) of Absolute Differences Mean (min,Q25,Q75,max)	Cases Mean / N, %	Controls Mean / N, %
Age (years)	0.13 (-3, 0, 0, 1, 3)	62.25	62.12
Enrollment date (days)	-1.93 (-75, -17, -3, 13, 69)	13445.5	13447.4
Baseline blood draw date	-1.14 (-112, -15, 1, 12, 139)	13387.3	13388.5
Year 3 blood draw date	0.74 (-154, -21, 2, 20, 137)	14617.8	14617.1
Ethnicity (1=white, 2=black, 3=hispanic, 4=other/unknown)	0	19 (27.5%) white 27 (39.1%) black 6 (8.7%) hispanic 17 (24.6%) other/unknown	
Hysterectomy at baseline	0	31 (44.9%) yes	
History of colorectal cancer (0=no, 1=yes)	0	5 (7.3%) yes	

Phase 7: WHI general population, matching baseline blood only

A total of 20 case / control pairs were successfully matched.

Matching Factor	Sum (weighted) of Absolute Differences Mean (min,Q25,Q75,max)	Cases Mean / N, %	Controls Mean / N, %
Age (years)	-0.25 (-2, -1, 0, 0, 2)	62.00	62.25
Enrollment date (days)	12.45 (-28, -5, 8, 19.5, 85)	13427.6	13415.1
Baseline blood draw date	5.80 (-43, -15.5, 3.5, 11.5, 84)	13385.2	13379.4
Ethnicity (1=white, 2=black, 3=hispanic, 4=other/unknown)	0	5 (25.0%) white 12 (60.0%) black 3 (15.0%) hispanic	
Hysterectomy at baseline	0	11 (55.0%) yes	
History of colorectal cancer (0=no, 1=yes)	0	4 (20.0%) yes	

Phase 8: WHI general population, relaxed matching parameters

A total of 5 case / control pairs were successfully matched.

Matching Factor	Sum (weighted) of Absolute Differences Mean (min,Q25,Q75,max)	Cases Mean / N, %	Controls Mean / N, %
Age (years)	-1.20 (-2, -2, -1, -1, 0)	65.20	66.40
Enrollment date (days)	-114.80 (-246, -172, -144, -139, 127)	12889.60	13004.4
Baseline blood draw date	41.60 (-233, -192, 14, 142, 477)	12826.2	12784.6
Year 3 blood draw date	-16.20 (-200, -196, -188, -58, 561)	14355.60	14371.8
Ethnicity (1=white, 2=black, 3=hispanic, 4=other/unknown)	0	2 (40.0%) white 3 (60.0%) black	
Hysterectomy at baseline	0	3 (60.0%) yes	
History of colorectal cancer (0=no, 1=yes)	0	2 (40.0%) yes	

Combined Phases:

A total of 988 case / control pairs were successfully matched.

Matching Factor	Sum (weighted) of Absolute Differences Mean (min,Q25,Q75,max)	Cases Mean / N, %	Controls Mean / N, %
Age (years)	-0.21 (-3, -1, 0, 0, 3)	66.63	66.84
Enrollment date (days)	3.72 (-317, -21, 0, 27, 316)	13465.0	13461.3
Baseline blood draw date (987 pair)	0.20 (-233, -34, 0, 32, 477)	13420.6	13420.4
Year 3 blood draw date (793 pair)	6.05 (-200, -29, 4, 41, 561)	14597.2	14591.1
Ethnicity (1=white, 2=black, 3=hispanic, 4=other/unknown)	0	842 (85.2%) white 88 (8.9%) black 22 (2.2%) hispanic 36 (3.6%) other/unknown	
Hysterectomy at baseline	0	416 (42.1%) yes	
History of colorectal cancer (0=no, 1=yes)	0	21 (2.1%) yes	