

Data as of September 20th, 2013

Population: CT/OS

Cases: All centrally adjudicated primary pancreatic adenocarcinoma (ICD-O-2 site code 25.0, 25.1, 25.2, 25.3, 25.4, 25.7, 25.8, 25.9) cases.

Controls: Participants who did not have a pancreatic cancer event as described above

Table 1. Sample size changes upon application of exclusion criteria

Exclusion Criteria	Aim 1 (Serum)		Aim 2 (DNA)	
	Cases	Controls	Cases	Controls
Original Sample	680	161128	680	161128
Exclude all participants w/ history of any cancer (except non-melanoma skin cancer)	623	146336	623	146336
Exclude all participants with missing matching variable data or no follow-up time	622	145617	622	145617
Exclude participants in any WHI HT OR DM active arm (Aim 1 only)	496	115289	-	-
Exclude all participants w/ inadequate serum (<250 ul, Aim 1 Only)	494	114310	-	-
Exclude all participants who have both inadequate DNA and are not members of an existing major GWAS ¹ (Aim 2 Only)	-	-	613	143599
Exclude all participants w/ a refusal for genetic studies (Aim 2 Only)	-	-	613	143583
Total post-exclusion participants	494	114310	613	143583

¹Major GWAS studies include GARNET, SHARE, WHIMS+, GECCO, BAA 03, and PanScan

A total of 494 cases and 114310 potential controls remained in the matching pool for aim 1, while 613 cases and 143583 potential controls remained in the matching pool for aim 2. There were a total of 486 cases and 112971 potential controls who were eligible in both matching pools.

Matching Phase 1: Participants eligible for both Aims 1 and 2

A total of 486 cases and 112971 potential controls were put into the matching process. Matching covariates included age (+/- 3 years), ethnicity, WHI randomization / enrollment date (+/- 6 months), blood draw date (+/- 6 months), OS enrollment (Yes/No), HT trial arm (E-Alone Placebo /E+P Placebo/Not Randomized), DM trial arm (Comparison/Not Randomized), CaD trial Enrollment (Yes/No), hysterectomy at baseline (Yes/No), and randomization clinic.

When selecting controls, a priority on selecting from participants enrolled in first SHARE and PANSCAN and then other major GWAS studies as well as a separate priority on participants with existing core analyte data was used. To account for this, additional matching criteria were added to the matching process.

Table 2. Final Phase 1 Matching Criteria, maximum distance, and weighting for matching

Criteria	Maximum Distance	Weighting
Age	3 years	45
Ethnicity	Exact	-
WHI randomization date	6 months	1
Blood draw date	6 months	1
OS enrollment	Exact	-
HT trial arm	Exact	-
DM trial arm	Exact	-
CaD trial enrollment	Exact	-
Hysterectomy	Exact	-
Randomization/Enrollment Clinic	Exact	-
Core analytes done (set to 'yes' for cases)	1	500
SHARE membership (set to 'yes' for cases)	1	500
PANSCAN membership (set to 'yes' for cases)	1	500
Major GWAS (set to 'yes' for cases)	1	500

Potential controls were selected based on criteria to minimize an overall distance measure, equating 1 year of age difference equivalent to a difference in 45 days in randomization date or blood draw date. Not being enrolled in SHARE, PANSCAN, any major GWAS, or not having core analyte data all equated with a difference of 500 days each in randomization date, thus prioritizing those potential controls who were members of these groups.

Of the 486 cases put into the matching algorithm, 429 were successfully matched with 2 controls, 17 matched only 1 control, and 40 were unable to match any controls.

For those cases that were unable to get two matches, the matching algorithm was then run a second time, this time removing the criteria to match clinical center and expanding the maximum age difference to 5 years.

For this second round of matching, all 17 cases who had one control in the previous match were able to match a second control. Of the 40 cases whom had no controls matched in the previous match, 38 were matched to 2 controls each, 1 was matched to 1 control, and 1 was unable to match any controls.

Matching was done in a time forward manner to ensure that each control had at as much control time as its matched case. For example, a case who developed pancreatic cancer two years after randomization would be matched with a control with at least two years of follow-up.

(Bergstralh EJ, Kosanke JL. Computerized matching of cases to controls. Technical Report #56, Department of Health Sciences Research, Mayo Clinic, Rochester, MN. April 1995).

Total participants from phase 1 include 485 cases and 969 controls for both Aims 1 (serum) and 2 (DNA).

Matching Phase 2: Participants eligible for Aim 1 only

A total of 9 cases and 113341 controls were put into the matching algorithm. Because these cases are ineligible for the genetic component, all genetic components of the matching were removed.

Table 3. Final Phase 2 Matching Criteria, maximum distance, and weighting for matching

Criteria	Maximum Distance	Weighting
Age	3 years	45
Ethnicity	Exact	-
WHI randomization date	6 months	1
Blood draw date	6 months	1
OS enrollment	Exact	-
HT trial arm	Exact	-
DM trial arm	Exact	-
CaD trial enrollment	Exact	-
Hysterectomy	Exact	-
Randomization/Enrollment Clinic	Exact	-
Core analytes done (set to 'yes' for cases)	1	500

Eight of the nine cases were able to match two controls. One was unable to match any controls. As in phase 1, for the case that was unable to get any matches, the matching algorithm was run a second time, removing the criteria to match clinical center and expanding the maximum age difference to 5 years.

For this second round of matching, one control was matched to the single case.

Total participants from phase 2 include 9 cases and 17 controls for Aim 1 (serum).

Matching Phase 3: Participants eligible for Aim 2 only

A total of 127 cases and 142599 controls were put into the matching algorithm. Because these cases are ineligible for the serum component, all serum components of the matching were removed.

Table 4. Final Phase 2 Matching Criteria, maximum distance, and weighting for matching

Criteria	Maximum Distance	Weighting
Age	3 years	45
Ethnicity	Exact	-
WHI randomization date	6 months	1
OS enrollment	Exact	-
HT trial arm	Exact	-
DM trial arm	Exact	-
CaD trial enrollment	Exact	-
Hysterectomy	Exact	-
Randomization/Enrollment Clinic	Exact	-
Core analytes done (set to 'yes' for cases)	1	500
SHARE membership (set to 'yes' for cases)	1	500
PANSCAN membership (set to 'yes' for cases)	1	500
Major GWAS (set to 'yes' for cases)	1	500

Out of a total of 127 cases put into the matching algorithm, 80 successfully matched 2 controls to each case, 17 were able to match 1 control to each case, and 30 were unable to match any controls. As in previous phases, for the cases that was unable to get two matches, the matching algorithm was run a second time, removing the criteria to match clinical center and expanding the maximum age difference to 5 years.

For this second round of matching, all 17 cases who had previously matched a control were able to match a second control. Of the 30 cases unable to match controls in the first round of matching, 1 was able to match 2 controls, 29 were unable to match any controls.

Total participants from phase 3 include 98 cases and 196 controls for Aim 2 (DNA).

Table 5. Matching Results – Aim 1 (Serum) Participants

Matching Factor	Sum (weighted) of Absolute Differences	Pancreatic Cancer Cases (n=494)	Pancreatic Cancer Controls (n=986)
	Mean (min, max)	Mean (SD) / n (%)	Mean (SD) / n (%)
Age	1.3 (0.0, 5.0)	65.4 (7.1)	65.3 (6.9)
Ethnicity	0		
White		430 (87.0)	860 (87.2)
African American		34 (6.9)	66 (6.7)
Hispanic		9 (1.8)	18 (1.8)
Native American		3 (0.6)	6 (0.6)
Asian / Pacific Islander		15 (3.0)	30 (3.0)
Unknown		3 (0.6)	6 (0.6)
WHI Randomization / Enrollment Date	54.3 (0.0, 182.0)	9/28/96 (401)	10/1/96 (397)
Blood draw date	54.9 (0.0, 182.0)	8/6/96 (413)	8/8/96 (410)
OS enrollment	0	327 (66.2)	654 (66.3)
HT Intervention Arm	0		
E-Alone Placebo		31 (6.3)	61 (6.2)
E+P Placebo		32 (6.5)	63 (6.4)
Not randomized		431 (87.3)	862 (87.4)
DM Intervention Arm	0		
Comparison		120 (24.3)	239 (24.2)
Not randomized		374 (75.7)	747 (75.8)
CaD Trial Enrollment	0	99 (20.0)	197 (20.0)
Hysterectomy	0	215 (43.5)	429 (43.5)
Randomization Clinic (% matched to case clinic)	-	-	875 (88.7)
Core analytes done	-	20 (4.1)	134 (13.6)

The average control follow-up time for cases was 8.6 years, for controls 14.8 years

Total Participants: 1480

Table 6. Matching Results – Aim 2 (DNA) Participants

Matching Factor	Sum (weighted) of Absolute Differences	Pancreatic Cancer Cases (n=583)	Pancreatic Cancer Controls (n=1165)
	Mean (min, max)	Mean (SD) / n (%)	Mean (SD) / n (%)
Age	1.3 (0.0, 5.0)	65.4 (7.0)	65.3 (6.8)
Ethnicity	0		
White		514 (88.2)	1028 (88.2)
African American		37 (6.4)	73 (6.3)
Hispanic		9 (1.5)	18 (1.6)
Native American		4 (0.7)	8 (0.7)
Asian / Pacific Islander		16 (2.7)	32 (2.8)
Unknown		3 (0.5)	6 (0.5)
WHI Randomization / Enrollment Date	57.8 (0.0, 182.0)	9/25/96 (405)	9/27/96 (400)
OS enrollment	0	321 (55.1)	642 (55.1)
HT Intervention Arm	0		
E-Alone Active		23 (4.0)	46 (4.0)
E-Alone Placebo		31 (5.3)	62 (5.3)
E+P Active		22 (3.8)	44 (3.8)
E+P Placebo		32 (5.5)	63 (5.4)
Not randomized		475 (81.5)	950 (81.6)
DM Intervention Arm	0		
Low-fat diet		52 (8.9)	104 (8.9)
Comparison		123 (21.1)	245 (21.0)
Not randomized		408 (70.0)	816 (70.0)
CaD Trial Enrollment	0	150 (25.7)	300 (25.8)
Hysterectomy	0	257 (44.1)	514 (44.1)
Randomization Clinic (% matched to case)	-	-	875 (75.1)
Core analytes done	-	25 (4.3)	152 (13.1)
SHARE membership	-	40 (6.9)	87 (7.5)
PanScan membership	-	444 (76.2)	50 (4.3)
GWAS ¹ Done	-	470 (80.6)	667 (57.3)

¹Major GWAS studies include GARNET, SHARE, WHIMS+, GECCO, BAA 03, and PanScan

The average control follow-up time for cases was 8.6 years, for controls 14.8 years

Total Participants: 1748

If participants who are already in a major GWAS do not need additional genetic testing, we have a total of 611 participants who need genetic analyses and all 1480 Aim 1 participants need serum analysis.