/\*\*\*\*\*\* Synthetic Controls Code \*\*\*\*\*\*/

USE Cancer

DROP TABLE #data, #cases

-- Cast variable characters variables as integers

SELECT [Column 0] AS ID

, CAST([Cancer] AS int) AS Cancer

, Cast([I305 1] AS Int) AS [I305\_1], Cast([I309 81] AS Int) AS [I309\_81]

, Cast([I311 ] AS Int) AS [I311], Cast([IE849 7] AS Int) AS [IE849\_7]

, Cast([I150 9] AS Int) AS [I150\_9], Cast([I276 1] AS Int) AS [I276\_1]

, Cast([I276 8] AS Int) AS [I276\_8], Cast([I530 81] AS Int) AS [I530\_81]

, Cast([I263 9] AS Int) AS [I263\_9], Cast([I276 51] AS Int) AS [I276\_51]

, Cast([IV15 82] AS Int) AS [IV15\_82], Cast([I511 9] AS Int) AS [I511\_9]

, Cast([I401 9] AS Int) AS [I401\_9], Cast([I787 20] AS Int) AS [I787\_20]

, Cast([I564 00] AS Int) AS [I564\_00], Cast([I272 4] AS Int) AS [I272\_4]

, Cast([I280 9] AS Int) AS [I280\_9], Cast([I285 9] AS Int) AS [I285\_9]

, Cast([I496 ] AS Int) AS [I496], Cast([I458 9] AS Int) AS [I458\_9]

, Cast([I486 ] AS Int) AS [I486], Cast([IV58 61] AS Int) AS [IV58\_61]

, Cast([I197 7] AS Int) AS [I197\_7], Cast([I578 9] AS Int) AS [I578\_9]

, Cast([I584 9] AS Int) AS [I584\_9], Cast([IV66 7] AS Int) AS [IV66\_7]

, Cast([I244 9] AS Int) AS [I244\_9], Cast([I414 01] AS Int) AS [I414\_01]

, Cast([I599 0] AS Int) AS [I599\_0], Cast([I414 00] AS Int) AS [I414\_00]

, Cast([I585 9] AS Int) AS [I585\_9], Cast([I600 00] AS Int) AS [I600\_00]

, Cast([I428 0] AS Int) AS [I428\_0], Cast([I427 31] AS Int) AS [I427\_31]

, Cast([I403 90] AS Int) AS [I403\_90], Cast([Dead] AS Int) AS [Dead]

INTO #Data

FROM [Cancer].[dbo].[StomachCancer]

-- (100000 row(s) affected)

-- Add regression predicted values to the data

Select \*

, exp(0.102056+I305\_1\*-0.022632+I309\_81\*-0.048045+I311\*-0.008337 +IE849\_7\*0.00235+I150\_9\*0.192467+I276\_1\*0.042832+I276\_8\*0.001069

+I530\_81\*-0.019579+I263\_9\*0.109362+I276\_51\*0.015216+IV15\_82\*-0.024939 +I511\_9\*0.074348+I401\_9\*0.004262+I787\_20\*0.08602+I564\_00\*0.033718

+I272\_4\*-0.030436+I280\_9\*0.025565+I285\_9\*0.065355+I496\*0.078868

+I458\_9\*0.04247+I486\*0.109481+IV58\_61\*-0.027556+I197\_7\*0.422058

+I578\_9\*0.076179+I584\_9\*0.056606+IV66\_7\*0.543452+I244\_9\*0.042195

+I414\_01\*0.006689+I599\_0\*0.077619+I414\_00\*0.045172+I585\_9\*0.00953

+I600\_00\*0.015642+I428\_0\*0.123643+I427\_31\*0.06699+I403\_90\*0.055357)

/(1+exp(0.102056+I305\_1\*-0.022632+I309\_81\*-0.048045+I311\*-0.008337+IE849\_7\*0.00235+I150\_9\*0.192467+I276\_1\*0.042832+I276\_8\*0.001069

+I530\_81\*-0.019579+I263\_9\*0.109362+I276\_51\*0.015216+IV15\_82\*-0.024939

+I511\_9\*0.074348+I401\_9\*0.004262+I787\_20\*0.08602+I564\_00\*0.033718

+I272\_4\*-0.030436+I280\_9\*0.025565+I285\_9\*0.065355+I496\*0.078868+I458\_9\*0.04247

+I486\*0.109481+IV58\_61\*-0.027556+I197\_7\*0.422058+I578\_9\*0.076179

+I584\_9\*0.056606+IV66\_7\*0.543452+I244\_9\*0.042195+I414\_01\*0.006689+I599\_0\*0.077619

+I414\_00\*0.045172+I585\_9\*0.00953+I600\_00\*0.015642+I428\_0\*0.123643+I427\_31\*0.06699

+I403\_90\*0.055357)) AS Predicted -- predicted probability of being dead

INTO #DATA1 FROM #DATA

-- Cases describe outcomes for cancer patients

DROP TABLE #Cases

SELECT COUNT(distinct [ID]) AS nCases

, Sum(IIF([Dead] = 1, 1., 0.)) AS a

, SUM(IIF([Dead] = 0, 1., 0.)) AS b

, Max([Predicted]) AS [Predicted Control] -- If case is not matched then predicted

, I305\_1, I309\_81, I311, IE849\_7, I150\_9, I276\_1, I276\_8, I530\_81, I263\_9, I276\_51, IV15\_82, I511\_9, I401\_9, I787\_20, I564\_00, I272\_4, I280\_9,

I285\_9, I496, I458\_9, I486, IV58\_61, I197\_7, I578\_9, I584\_9, IV66\_7, I244\_9, I414\_01, I599\_0, I414\_00, I585\_9, I600\_00, I428\_0, I427\_31, I403\_90

INTO #Cases -- Save in temporary file called Case

FROM #Data1

WHERE [Cancer]=1

GROUP BY

I305\_1, I309\_81, I311, IE849\_7, I150\_9, I276\_1, I276\_8, I530\_81, I263\_9, I276\_51, IV15\_82, I511\_9, I401\_9, I787\_20, I564\_00, I272\_4, I280\_9,

I285\_9, I496, I458\_9, I486, IV58\_61, I197\_7, I578\_9, I584\_9, IV66\_7, I244\_9, I414\_01, I599\_0, I414\_00, I585\_9, I600\_00, I428\_0, I427\_31, I403\_90

-- (180 row(s) affected)

-- Controls describe patients without cancer

drop table #Controls

SELECT COUNT(distinct [ID]) AS nControls

, Sum(IIF([Dead] = 1, 1., 0.)) AS c

, SUM(IIF([Dead] = 0, 1., 0.)) AS d

,I305\_1, I309\_81, I311, IE849\_7, I150\_9, I276\_1, I276\_8, I530\_81, I263\_9, I276\_51, IV15\_82, I511\_9, I401\_9, I787\_20, I564\_00, I272\_4, I280\_9,

I285\_9, I496, I458\_9, I486, IV58\_61, I197\_7, I578\_9, I584\_9, IV66\_7, I244\_9, I414\_01, I599\_0, I414\_00, I585\_9, I600\_00, I428\_0, I427\_31, I403\_90

INTO #Controls -- Save in temporary file called Controls

FROM #Data

WHERE [Cancer]=0

GROUP BY I305\_1, I309\_81, I311, IE849\_7, I150\_9, I276\_1, I276\_8, I530\_81, I263\_9, I276\_51, IV15\_82, I511\_9, I401\_9, I787\_20, I564\_00, I272\_4, I280\_9,

I285\_9, I496, I458\_9, I486, IV58\_61, I197\_7, I578\_9, I584\_9, IV66\_7, I244\_9, I414\_01, I599\_0, I414\_00, I585\_9, I600\_00, I428\_0, I427\_31, I403\_90

-- (26761 row(s) affected)

-- Match cases with controls and calculate odds ratio and percent overlap

DROP TABLE #Match

SELECT a

,b

,iif(c is null, (a+b)\*[Predicted Control], c) AS c -- Replaces missing dead controls with predicted counts

,iif(d is null, (a+b)\*(1-[Predicted Control]),d) AS d -- Replaces missing alive controls with 1-predicted counts

INTO #Match

FROM #Cases left join #Controls -- left join as cases not matched must have a null control

ON #Cases.[I305\_1] =#Controls.[I305\_1] and #Cases.[I309\_81] = #Controls.[I309\_81] and #Cases.[I311]= #Controls.[I311] and #Cases.[IE849\_7]= #Controls.[IE849\_7] and #Cases.[I150\_9]= #Controls.[I150\_9] and #Cases.[I276\_1]= #Controls.[I276\_1]

and #Cases.[I276\_8]= #Controls.[I276\_8] and #Cases.[I530\_81]= #Controls.[I530\_81] and #Cases.[I263\_9]= #Controls.[I263\_9] and #Cases.[I276\_51]= #Controls.[I276\_51] and #Cases.[IV15\_82]= #Controls.[IV15\_82] and #Cases.[I511\_9]= #Controls.[I511\_9]

and #Cases.[I401\_9]= #Controls.[I401\_9] and #Cases.[I787\_20]= #Controls.[I787\_20] and #Cases.[I564\_00]= #Controls.[I564\_00] and #Cases.[I272\_4]= #Controls.[I272\_4] and #Cases.[I280\_9]= #Controls.[I280\_9] and #Cases.[I285\_9]= #Controls.[I285\_9]

and #Cases.[I496]= #Controls.[I496] and #Cases.[I458\_9]= #Controls.[I458\_9] and #Cases.[I486]= #Controls.[I486] and #Cases.[IV58\_61]= #Controls.[IV58\_61] and #Cases.[I197\_7]= #Controls.[I197\_7] and #Cases.[I578\_9]= #Controls.[I578\_9]

and #Cases.[I584\_9]= #Controls.[I584\_9] and #Cases.[IV66\_7]= #Controls.[IV66\_7] and #Cases.[I244\_9]= #Controls.[I244\_9] and #Cases.[I414\_01]= #Controls.[I414\_01] and #Cases.[I599\_0]= #Controls.[I599\_0] and #Cases.[I414\_00]= #Controls.[I414\_00]

and #Cases.[I585\_9]= #Controls.[I585\_9] and #Cases.[I600\_00]= #Controls.[I600\_00] and #Cases.[I428\_0]= #Controls.[I428\_0] and #Cases.[I427\_31]= #Controls.[I427\_31] and #Cases.[I403\_90]= #Controls.[I403\_90]

-- Calculate Odds Ratio & Overlap

SELECT \* FROM #Match

SELECT

sum(a\*d/(a+b+c+d))/sum(b\*c/(a+b+c+d)) As [Common Odds Ratio]

, ROUND(100.\*SUM(iif(c is not null, a+b, 0))/sum(a+b),2) as [Percent Overlap]

FROM #Match