**Appendix 1**

PolyMAIC SAS code for Scenario A (narrow tolerances):

\*\* ALD targets & matching tolerances \*\*;

data ald\_targ;

%let m1=55; \* Mean age;

%let s1=6; \* SD age;

%let mto1=0.005; \* Mean age – matching tolerance;

%let sto1=0.005; \* SD age – matching tolerance;

%let m2=0.65; \* Proportion female;

%let mto2=0.0005; \* Proportion female – matching tolerance;

run;

PROC OPTMODEL;

set OBS;\* Number of IPD observations (N);

\* Read in IPD.

x1 = min-max ‘age’ scaling.

x2 = min-max ‘female proportion’ scaling;

num AGE{OBS},SEXF{OBS},x1{OBS},x2{OBS};

read data IPD1 into OBS = [i] AGE SEXF x1 x2;

var b1 {j1 in 0..4}; \* 4th order polynomial parameters for age;

var b2 {j1 in 1..1}; \* 1st order polynomial parameter for sex;

\* Weight calculations;

impvar w1 {i in OBS} = b1[0]+b1[1]\*x1[i]+b1[2]\*x1[i]\*\*2+b1[3]\*x1[i]\*\*3+b1[4]\*x1[i]\*\*4; \* Age;

impvar w2 {i in OBS} = b2[1]\*x2[i]; \* Proportion female;

impvar w {i in OBS} = exp(w1[i]+w2[i]); \* Exp. combined weights;

impvar sw = sum {i in OBS} w[i]; \* Sum of exp. combined weights;

\* weighted means;

impvar wm1 = sum {i in OBS} AGE[i] \*w[i] / sw; \* age;

impvar wm2 = sum {i in OBS} SEXF[i]\*w[i] / sw; \* prop. female;

\* weighted SDs (continuous characteristic only);

impvar wsd1 = sqrt(sum {i in OBS} w[i]\*(AGE[i]-wm1)\*\*2/(sw-1)); \* age;

\* Effective sample size (ESS) and max. weight;

impvar ess = sw\*\*2 / sum {i in OBS} w[i]\*\*2; \* ESS;

impvar maxw = max {i in OBS} w[i]; \* Maximum observed weight;

\* Constraints;

con r1: sum {i in OBS} w[i] = card(OBS); \* Weights sum to N;

\* Constraints (weighted IPD vs. ALD tolerance within target);

con r2: abs(&m1-wm1) <= &mto1; \* Mean age;

con r3: abs(&s1-wsd1) <= &sto1; \* SD age;

con r4: abs(&m2-wm2) <= &mto2; \* Proportion female;

\*con r5: max {i in OBS} w[i] <= 5; \* Max. weight restriction (optional);

\* Find solution;

max objective = ess; \* Maximise ESS according to above constraints;

solve with nlp / multistart=(maxstarts=20) seed=43113; \*Non-linear optimization;

\* Output matching performance;

print wm1 wsd1 wm2 ess maxw;

\* Output IPWs;

create data ad.poly11 from [OBS] AGE SEXF w; \* weights;

create data ad.poly11p from [OBS] b1 b2; \* parameters for polynomial;

quit;

PolyMAIC SAS code for Scenario D (wide tolerances):

\*\* ALD targets & matching tolerances \*\*;

data t1;

%let m1=57.0; \* Mean age;

%let s1=7.2; \* SD age;

%let mto1=2; \* Mean age - matching tolerance;

%let sto1=0.5; \* SD age - matching tolerance;

%let m2=0.568; \* Proportion female;

%let mto2=0.02; \* Proportion female - matching tolerance;

%let m3=7.2; \* Mean pain;

%let s3=1.3; \* SD pain;

%let mto3=1; \* Mean pain - matching tolerance;

%let sto3=0.5; \* SD pain - matching tolerance;

%let m4=12.3; \* Mean disease duration;

%let s4=6.4; \* SD disease duration;

%let mto4=1; \* Mean disease duration - matching tolerance;

%let sto4=1; \* SD disease duration - matching tolerance;

%let m5=1.8; \* Mean # prior events;

%let s5=1.9; \* SD # prior events;

%let mto5=0.25; \* Mean # prior events - matching tolerance;

%let sto5=0.25; \* SD # prior events - matching tolerance;

run;

proc optmodel;

set OBS; \* Number of IPD observations (N);

\* Read in IPD.

x1 = min-max ‘age’ scaling

x2 = min-max ‘female proportion’ scaling

x3 = min-max 'pain' scaling

x4 = min-max 'disease duration' scaling

x5 = min=max '# prior events' scaling;

num AGE{OBS}, SEXF{OBS}, PAIN{OBS}, TIMED{OBS}, EPIS{OBS}, x1{OBS}, x2{OBS},x3{OBS},x4{OBS},x5{OBS};

read data d1 into OBS = [i] AGE SEXF PAIN TIMED EPIS x1 x2 x3 x4 x5;

var a1 {j1 in 0..4}; \* 4th order polynomial parameters for age;

var a2 {j1 in 1..1}; \* 1st order polynomial parameter for sex;

var a3 {j1 in 1..4}; \* 4th order polynomial parameters for pain;

var a4 {j1 in 1..4}; \* 4th order polynomial parameters for disease duration;

var a5 {j1 in 1..4}; \* 4th order polynomial parameters for #prior events;

\* Weight calculations;

impvar w1 {i in OBS} = a1[0]+a1[1]\*x1[i]+a1[2]\*x1[i]\*\*2+a1[3]\*x1[i]\*\*3+a1[4]\*x1[i]\*\*4; \* age;

impvar w2 {I in OBS} = a2[1]\*x2[i]; \* proportion female;

impvar w3 {i in OBS} = a3[1]\*x3[i]+a3[2]\*x3[i]\*\*2+a3[3]\*x3[i]\*\*3+a3[4]\*x3[i]\*\*4; \* pain;

impvar w4 {i in OBS} = a4[1]\*x4[i]+a4[2]\*x4[i]\*\*2+a4[3]\*x4[i]\*\*3+a4[4]\*x4[i]\*\*4; \* disease duration;

impvar w5 {i in OBS} = a5[1]\*x5[i]+a5[2]\*x5[i]\*\*2+a5[3]\*x5[i]\*\*3+a5[4]\*x5[i]\*\*4; \* # prior events;

impvar w {i in OBS} = exp(w1[i]+w2[i]+w3[i]+w4[i]+w5[i]); \* Exp. combined weights;

impvar sw = sum {i in OBS} w[i]; \* Sum of exp. combined weights;

\* weighted means;

impvar wm1 = sum {i in OBS} AGE[i]\*w[i] / sw; \* age;

impvar wm2 = sum {i in OBS} SEXF[i]\*w[i] / sw; \* prop. female;

impvar wm3 = sum {I in OBS} PAIN[i]\*w[i] / sw; \* pain;

impvar wm4 = sum {I in OBS} TIMED[i]\*w[i] / sw; \* disease duration;

impvar wm5 = sum {i in OBS} EPIS[i]\*w[i] / sw; \* # prior events;

\* weighted SDs (continuous characteristics only);

impvar wsd1 = sqrt(sum {I in OBS} w[i]\*(AGE[i]-wm1)\*\*2/(sw-1)); \* age;

impvar wsd3 = sqrt(sum {i in OBS} w[i]\*(PAIN[i]-wm3)\*\*2/(sw-1)); \* pain;

impvar wsd4 = sqrt(sum {i in OBS} w[i]\*(TIMED[i]-wm4)\*\*2/(sw-1)); \* disease duration;

impvar wsd5 = sqrt(sum {i in OBS} w[i]\*(EPIS[i]-wm5)\*\*2/(sw-1)); \* # prior events;

\* Effective sample size (ESS) and max. weight;

impvar ess = sw\*\*2 / sum {i in OBS} w[i]\*\*2; \* ESS;

impvar maxw = max {I in OBS} w[i]; \* Maximum observed weight;

\* Constraints;

con r1: sum {i in OBS} w[i] = card(OBS); \* Weights sum to N;

\* Constraints (weighted IPD vs. ALD tolerance within target);

con r2: abs(&m1-wm1)<=&mto1; \* Mean age;

con r3: abs(&s1-wsd1)<=&sto1; \* SD age;

con r4: abs(&m2-wm2)<=&mto2; \* Pop. female;

con r5: abs(&m3-wm3)<=&mto3; \* Mean pain;

con r6: abs(&s3-wsd3)<=&sto3; \* SD pain;

con r7: abs(&m4-wm4)<=&mto4; \* Mean disease duration;

con r8: abs(&s4-wsd4)<=&sto4; \* SD disease duration;

con r9: abs(&m5-wm5)<=&mto5; \* Mean # prior events;

con r10: abs(&s5-wsd5)<=&sto5; \*SD # prior events;

con r11: max {i in OBS} w[i] <=10; \* Max. weight restriction (optional);

\* Find solution;

max objective = ess; \* Maximise ESS according to above constraints;

solve with nlp / multistart=(maxstarts=50) seed=31211; \* non-linear optimization;

\* Output matching performance;

print wm1 wsd1 wm2 wm3 wsd3 wm4 wsd4 wm5 wsd5 ess maxw;

\* Output IPWs;

create data ad.poly42 from [OBS] AGE SEXF PAIN TIMED EPIS w; \* weights;

create data ad.poly42p from [OBS] a1 a2 a3 a4 a5; \* parameters for polynomial;

quit;