**Supplemental Digital Content 4. SAS code for the simulations.**

/\*11/2017\*/

/\*Analyst: Kristin Sainani

/\*A re-analysis of type I and type II errors in "Magnitude-Based Inference"\*/

/\*This re-uses the code as published in the supplement of: Hopkins WG, Batterham AM. Error Rates, Decisive Outcomes and Publication Bias with Several Inferential Methods. Sports Med. 2016 Oct;46(10):1563-73.\*/

/\*I have modified the code to systematically modify the sample size, true effect size, and minimally clinically significant threshold (StadzedMagniThresh).

/\*I have also taken the code out of Macro form.\*/

/\*I have also commented out some pieces of code that were not needed for my final simulations and graphics\*/

/\*For the calculation of type I and type II errors, I am:

1. Using Hopkins/Batterham's definitions without modification = defs1.

2. Correcting Hopkins/Batterham's definitions for clinical MBI = defsclinical

3. Correcting Hopkins/Batterham’s definitions for nonclinical MBI = defsnonc

/\*My comments are noted with KS\*/

/\*KS: This is an addition to improve the speed of my simulations\*/

**%macro** ODSOff(); /\* Call prior to BY-group processing \*/

ods graphics off;

ods exclude all;

ods noresults;

**%mend**;

**%macro** ODSOn(); /\* Call after BY-group processing \*/

ods graphics on;

ods exclude none;

ods results;

**%mend**;

/\*KS: Set the parameters of the simulation\*/

%LET LogFlag=0;

%LET NoOfTrials=100000;

\*%LET NoA=50;

\*%LET NoB=50;

\*%LET ES=0.2;

%LET Mean=0;

%LET StdDev=1;

%LET r=0.818; \*KS: Using same value as in Sports Med 2016;

\*%LET StdzedMagniThresh=.2;

/\*KS: Generate the simulated data\*/

/\*KS: I generated datasets with effect sizes of -0.3 to +0.3 across a range of different sample sizes (10 to 150 by 10).

/\*KS: I also generated a dataset where the effect size is fixed at 0, but I vary the smallest important effect (trivial threshold) from 0.1 to 0.3 across a wider range of sample sizes (5 to 300 by 5). The wider range of sample sizes is necessary to fully see the pattern.\*/

**data** dat;

do No=**5** to **300** by **5**; \*KS:Allows systematic variation of parameters;

do StdzedMagniThresh=**.1** to **.3**;

do ES=**0** to **0** by **1**;

do Trial=**1** to &NoOfTrials;

Group="A";

do athlete=**1** to No;

Y=&Mean+&StdDev\*rannor(**0**);

Y1=Y+sqrt(**1**/&r-**1**)\*&StdDev\*rannor(**0**);

Y2=Y+sqrt(**1**/&r-**1**)\*&StdDev\*rannor(**0**);

DeltaY=Y2-Y1;

output;

end;

Group="B";

do athlete=No+**1** to **2**\*No;

Y=&Mean+&StdDev\*rannor(**0**);

Y1=Y+sqrt(**1**/&r-**1**)\*&StdDev\*rannor(**0**);

Y2=Y+sqrt(**1**/&r-**1**)\*&StdDev\*rannor(**0**)+ES\*&StdDev/sqrt(&r);\*using observed SD to standardize;

DeltaY=Y2-Y1;

output;

end;

end;

end;

end;

end;

**run**;

/\*

\*check things;

proc means data=dat maxdec=1 fw=7;

class trial group;

var Y Y1 Y2 DeltaY;

run;

\*/

/\*data clev pred pred1 pred2 est est1 estCohen lsm lsm1 lsmdif lsmdif1 difCohen

cov cov0 cov1 cov2 cov3 covsum solf solf1 solr solr1;

ods listing close;\*/

/\*KS: The next set of code generates the statistics for each simulated dataset\*/

/\*KS: The ODSOff Macro speeds up the simulation\*/

%***ODSOff***

options nonotes; /\* KS: use NONOTES to suppress notes to the log \*/

**proc** **mixed** data=dat covtest cl alpha=**0.1**;

class Group ;

model DeltaY=Group/ddfm=sat;

\*lsmeans Group/diff=control('Control') alpha=0.1;

estimate "Mean change" group -**1** **1**/cl alpha=**0.1**;

repeated/group=Group; \*estimates different SD of change scores in the groups;

ods output estimates=est;

by No StdzedMagniThresh ES Trial;

/\*

ods output classlevels=clev;

ods output lsmeans=lsm;

ods output diffs=lsmdiff;

ods output covparms=cov;

ods output solutionf=solf;

ods output solutionr=solr;

\*/

**run**;

ods listing;

options notes; /\*KS: turn NOTES back on \*/

%***ODSOn***

\*title2 "Baseline SD for standardizing";

\*title3 "by averaging the variances in the two groups";

**proc** **means** noprint data=dat;

var Y1;

by No StdzedMagniThresh ES Trial;

\*by group;

output out=stdsd std=PreSD n=NoOfObs;

/\*

proc print data=stdsd;

where Trial<100;

run;

proc means data=stdsd;

var preSD;

run;

\*/

\*title2 "Standardized fixed effects";

**data** est1;

merge est stdsd(keep=No StdzedMagniThresh ES Trial PreSD NoOfObs);

by No StdzedMagniThresh ES Trial;

\*if estimate=0 then estimate=.;

array a estimate lower upper;

do over a;

a=a/PreSD\*(**1**-**3**/(**4**\*(NoOfObs-**1**)-**1**)); \*correction for bias in stdzd diff in means;

end;

CLpm=(Upper-lower)/**2**;

/\*

proc print data=est1 noobs;

var Trial Label estimate CLpm lower upper alpha DF Probt;

format estimate stderr CLpm lower upper 6.2 Probt best5. DF 5.0;

run;

\*/

/\*KS: The dataset est1 will contain the MBI and standard hypothesis testing results for the simulated trials\*/

\*title2 "MBI for standardized fixed effects";

data est1;

merge est stdsd(keep=No StdzedMagniThresh ES Trial PreSD NoOfObs);

by No StdzedMagniThresh ES Trial;

LCL99=estimate+stderr\*tinv(**.005**,df);

LCL90=estimate+stderr\*tinv(**.05**,df);

LCL50=estimate+stderr\*tinv(**.25**,df);

UCL50=estimate+stderr\*tinv(**.75**,df);

UCL90=estimate+stderr\*tinv(**.95**,df);

UCL99=estimate+stderr\*tinv(**.995**,df);

array a estimate lower upper stderr LCL99--UCL99;

do over a;

a=a/PreSD\*(**1**-**3**/(**4**\*(NoOfObs-**1**)-**1**));

end;

\*StdzedMagniThresh=&StdzedMagniThresh; \*KS: I set this parameter earlier to make it easier to vary the parameter systematically;

if &LogFlag=**1** then do;

EquivPcentThresh=**100**\*exp(StdzedMagniThresh\*PreSD/**100**)-**100**;

if StdzedMagniThresh>**0** then do;

ChancePos=**100**\*(**1**-ProbT(-(estimate-**100**\*log(**1**+StdzedMagniThresh/**100**))/StdErr,DF));

ChanceNeg=**100**\*ProbT(-(estimate+**100**\*log(**1**+StdzedMagniThresh/**100**))/StdErr,DF);

end;

else do;

ChancePos=**100**\*(**1**-ProbT(-(estimate+**100**\*log(**1**+StdzedMagniThresh/**100**))/StdErr,DF));

ChanceNeg=**100**\*ProbT(-(estimate-**100**\*log(**1**+StdzedMagniThresh/**100**))/StdErr,DF);

end;

end;

else do;

EquivRawThresh=StdzedMagniThresh\*PreSD;

ChancePos=**100**\*(**1**-ProbT(-(estimate-abs(StdzedMagniThresh))/StdErr,DF));

ChanceNeg=**100**\*ProbT(-(estimate+abs(StdzedMagniThresh))/StdErr,DF);

end;

ChanceTriv=**100**-ChancePos-ChanceNeg;

ORPosNeg=ChancePos/(**100**-ChancePos)/(ChanceNeg/(**100**-ChanceNeg));

ORNegPos=**1**/ORPosNeg;

ClinFlag=**1**; \*want inferences to be clinical;

if index(label,"2SD") then ClinFlag=**0**; \*covariates definitely need to be non-clinical;

\*clinical inferences;

if clinflag then do;

ClearOrNot="unclear";

ChPos=ChancePos; ChNeg=ChanceNeg;

if StdzedMagniThresh<**0** then do;

ChPos=ChanceNeg; ChNeg=ChancePos;

end;

if ChNeg<**0.5** and ChPos>**25** then ClearOrNot="@25/.5%"; \*not harm at the 0.5;

if ChNeg<**0.1** and ChPos>**25** then ClearOrNot="@5/.1% "; \*not harm at the 0.1;

if ClearOrNot="unclear" and (StdzedMagniThresh>**0** and ORPosNeg>**25**/**75**/(**0.5**/**99.5**)

or StdzedMagniThresh<**0** and ORNegPos>**25**/**75**/(**0.5**/**99.5**))

then ClearOrNot="OR>66.3";

if ClearOrNot ne "unclear" then do; \*must be some kind of beneficial at this point;

Magni="3.bene";

p=ChPos;

end;

if ClearOrNot="unclear" then do; \*sort out if clearly trivial or harmful;

if ChPos<**25** then ClearOrNot="@25/.5%"; \*not bene at the 25%;

if ChPos<**5** then ClearOrNot="@5/.1% "; \*not bene at the 5%;

if ClearOrNot ne "unclear" then do;

p=ChanceTriv; Magni="2.triv";

if ChNeg>**75** then do;

p=ChNeg; Magni="1.harm";

end;

end;

end;

if p=**.** then Prob="0.unclear ";

if p>**.** then Prob="1.m.unlikely";

if p>**0.5** then Prob="2.v.unlikely";

if P>**5** then Prob="3.unlikely ";

if P>**25** then Prob="4.possibly ";

if P>**75** then Prob="5.likely ";

if P>**95** then Prob="6.v.likely ";

if P>**99.5** then Prob="7.m.likely ";

output;

end;

\*mechanistic inferences;

ClinFlag=**0**;

ClearOrNot="unclear";

Prob="";

Magni="";

ORPosNeg=**.**; ORNegPos=**.**;

if ChanceNeg<**5** or ChancePos<**5** then ClearOrNot="@90% ";

if ChanceNeg<**0.5** or ChancePos<**0.5** then ClearOrNot="@99% ";

Prob="0.unclear ";

if ClearOrNot ne "unclear" then do;

Magni="3.+ive";

if estimate<**0** then Magni="1.-ive";

if ChancePos>**5** or ChanceNeg>**5** then Prob="3.unlikely";

if ChancePos>**25** or ChanceNeg>**25** then Prob="4.possibly";

if ChancePos>**75** or ChanceNeg>**75** then Prob="5.likely ";

if ChancePos>**95** or ChanceNeg>**95** then Prob="6.v.likely";

if ChancePos>**99.5** or ChanceNeg>**99.5** then Prob="7.m.likely";

end;

if ClearOrNot ne "unclear" and ChanceTriv>**75** then do;

Magni="2.triv";

Prob="5.likely ";

if ChanceTriv>**95** then Prob="6.v.likely";

if ChanceTriv>**99.5** then Prob="7.m.likely";

end;

\*end;

output;

**run**;

/\*proc freq data=est1;

tables prob\*magni/norow nocol nofreq missing;

where clinflag=1;

\*by StdzdEffect;

run;\*/

/\*KS: Defs1 contains Hopkins and Batterham’s definitions\*/

/\*KS: I have commented out definitions that I didn’t explore (for strict NHST and OR cases)\*/

**data** defs1;

set est1;

\*data lsmdiff2;

\*set lsmdiff1;

if estimate=**0** then do;

estimate=**.**; StdzedMagniThresh=**.**; EquivRawThresh=**.**; EquivPcentThresh=**.**; magni=""; clearornot="";

end;

rename df=DegFree;

CLpm=(Upper-lower)/**2**;

ClinMeanClear=**.**;

ClinType1a=**0**;

ClinType1b=**0**;

ClinType2a=**0**;

ClinType2b=**0**;

ClinTypeIa=**0**;

ClinTypeIb=**0**;

ClinType1=**0**;

ClinType2=**0**;

ClinTypeI=**0**;

ClinTypeII=**0**;

ClinUnclear=**0**;

if LCL99<-StdzedMagniThresh and UCL50>StdzedMagniThresh then ClinUnclear=**100**;

if ClinUnclear=**0** then do; \*do all the following only for clear effects;

if ES>=StdzedMagniThresh then do; \*true effect is beneficial;

if -StdzedMagniThresh<LCL99 and UCL50<StdzedMagniThresh then ClinType2a=**100**;

if LCL99<-StdzedMagniThresh and -StdzedMagniThresh<UCL50<StdzedMagniThresh then ClinType2b=**100**;

if UCL50<-StdzedMagniThresh then ClinType2b=**100**;

end;

if -StdzedMagniThresh<ES<StdzedMagniThresh then do; \*true effect is trivial;

if StdzedMagniThresh<UCL50 then ClinTypeIb=**100**; \*possibly bene;

if UCL90<-StdzedMagniThresh then ClinTypeIa=**100**; \*v.unlikely trivial on harm side;

end;

if ES<=-StdzedMagniThresh then do; \*true effect is harmful;

if StdzedMagniThresh<LCL99 then ClinType1b=**100**;

if -StdzedMagniThresh<LCL99<StdzedMagniThresh and StdzedMagniThresh<UCL50 then ClinType1b=**100**;

if -StdzedMagniThresh<LCL99 and UCL50<StdzedMagniThresh then ClinType1a=**100**;

end;

ClinMeanClear=estimate;

ClinType1=ClinType1a+ClinType1b; \*failing to find a harmful effect when you should;

ClinType2=ClinType2a+ClinType2b; \*correct type II error;

ClinTypeI=ClinTypeIa+ClinTypeIb;\*correct type I error;

ClinTypeII=ClinType1+ClinType2;

end;

\*for OR approach, same as above, but have to fix when OR>66.3 and (obsvd=unclear or wholly trivial);

/\*OddRatMeanClear=.;

OddRatType1a=ClinType1a;

OddRatType1b=ClinType1b;

OddRatType2a=ClinType2a;

OddRatType2b=ClinType2b;

OddRatTypeIa=ClinTypeIa;

OddRatTypeIb=ClinTypeIb;

OddRatType1=0;

OddRatType2=0;

OddRatTypeI=0;

OddRatTypeII=0;

OddRatUnclear=ClinUnclear;

if ORPosNeg>25/75/(0.5/99.5) and ClinUnclear=100 then do;

OddRatUnclear=0; \*reset this;

if -&StdzedMagniThresh<&ES<&StdzedMagniThresh then OddRatTypeIb=100; \*if true=triv, error=TypeIb;

if &ES<=-&StdzedMagniThresh then OddRatType1b=100; \*if true=harm, error=Type1b;

end;

if ORPosNeg>25/75/(0.5/99.5) and -&StdzedMagniThresh<LCL99 and UCL50<&StdzedMagniThresh then do; \*all trivial;

if &ES>=&StdzedMagniThresh then OddRatType2a=0; \*if true=bene, reset this error;

if -&StdzedMagniThresh<&ES<&StdzedMagniThresh then OddRatTypeIb=100; \*if true=triv, error=TypeI;

if &ES<=-&StdzedMagniThresh then do; \*if true=harm;

OddRatType1b=100; \* error=Type1b;

OddRatType1a=0; \*and reset this error;

end;

end;

if OddRatUnclear=0 then OddRatMeanClear=estimate;

OddRatType1=OddRatType1a+OddRatType1b;

OddRatType2=OddRatType2a+OddRatType2b;

OddRatTypeI=OddRatTypeIa+OddRatTypeIb;

OddRatTypeII=OddRatType1+OddRatType2;

NoncMeanClear=**.**;

NoncType1a=**0**;

NoncType1b=**0**;

NoncType1=**0**;

NoncTypeI=**0**;

NoncTypeII=**0**;

NoncUnclear=**0**;

if LCL90<-StdzedMagniThresh and UCL90>StdzedMagniThresh then NoncUnclear=**100**;

if NoncUnclear=**0** then do; \*do all the following only for clear effects;

if ES>=StdzedMagniThresh then do; \*true effect is positive;

if -StdzedMagniThresh<LCL90 and UCL90<StdzedMagniThresh then NoncType1a=**100**;

if LCL90<-StdzedMagniThresh and -StdzedMagniThresh<UCL90<StdzedMagniThresh then NoncType1a=**100**;

if UCL90<-StdzedMagniThresh then NoncType1b=**100**;

end;

if -StdzedMagniThresh<ES<StdzedMagniThresh then do; \*true effect is trivial;

if StdzedMagniThresh<LCL90 then NoncTypeI=**100**;

if UCL90<-StdzedMagniThresh then NoncTypeI=**100**;

end;

if ES<=-StdzedMagniThresh then do; \*true effect is negative;

if StdzedMagniThresh<LCL90 then NoncType1b=**100**;

if -StdzedMagniThresh<LCL90<StdzedMagniThresh and StdzedMagniThresh<UCL90 then NoncType1a=**100**;

if -StdzedMagniThresh<LCL90 and UCL90<StdzedMagniThresh then NoncType1a=**100**;

end;

NoncMeanClear=estimate;

NoncType1=NoncType1a+NoncType1b;

NoncTypeII=NoncType1;

end;

NHpopClinMeanAll=**.**;

NHpopClinMeanSig=**.**;

NHpopClinMeanSigBene=**.**;

NHpopClinType1a=**0**;

NHpopClinType1b=**0**;

NHpopClinTypeIa=**0**;

NHpopClinTypeIb=**0**;

NHpopClinTypeIIa=**0**;

NHpopClinTypeIIb=**0**;

NHpopClinType1=**0**;

NHpopClinTypeI=**0**;

NHpopClinTypeII=**0**;

NHpopClinNonsig=**0**;

if Probt>**0.05** then NHpopClinNonsig=**100**;

if ES>=StdzedMagniThresh then do; \*true effect is beneficial;

if Probt>**0.05** then NHpopClinTypeIIa=**100**;

if Probt<**0.05** and estimate<**0** then NHpopClinTypeIIb=**100**;

end;

if -StdzedMagniThresh<ES<StdzedMagniThresh then do; \*true effect is trivial;

if Probt<**0.05** and estimate>**0** then NHpopClinTypeIb=**100**;

if Probt<**0.05** and estimate<**0** then NHpopClinTypeIa=**100**;

end;

if ES<=-StdzedMagniThresh then do; \*true effect is harmful;

if Probt<**0.05** and estimate>**0** then NHpopClinType1b=**100**;

if Probt>**0.05** then NHpopClinType1a=**100**;

end;

NHpopClinMeanAll=estimate;

if Probt<**0.05** then NHpopClinMeanSig=estimate;

if Probt<**0.05** and estimate>**0** then NHpopClinMeanSigBene=estimate;

NHpopClinType1=NHpopClinType1a+NHpopClinType1b;

NHpopClinTypeI=NHpopClinTypeIa+NHpopClinTypeIb;

NHpopClinTypeII=NHpopClinTypeIIa+NHpopClinTypeIIb;

NHpopClinTypeII=NHpopClinTypeII+NHpopClinType1;

NHpopNoncMeanAll=**.**;

NHpopNoncMeanSig=**.**;

NHpopNoncMeanSigPos=**.**;

NHpopNoncTypeIIa=**0**;

NHpopNoncTypeIIb=**0**;

NHpopNoncTypeI=**0**;

NHpopNoncTypeII=**0**;

NHpopNoncNonsig=**0**;

if Probt>**0.05** then NHpopNoncNonsig=**100**;

if ES>=StdzedMagniThresh then do; \*true effect is positive;

if Probt>**0.05** then NHpopNoncTypeIIa=**100**;

if Probt<**0.05** and estimate<**0** then NHpopNoncTypeIIb=**100**;

end;

if -StdzedMagniThresh<ES<StdzedMagniThresh then do; \*true effect is trivial;

if Probt<**0.05** and estimate>**0** then NHpopNoncTypeI=**100**;

if Probt<**0.05** and estimate<**0** then NHpopNoncTypeI=**100**;

end;

if ES<=-StdzedMagniThresh then do; \*true effect is negative;

if Probt<**0.05** and estimate>**0** then NHpopNoncTypeIIb=**100**;

if Probt>**0.05** then NHpopNoncTypeIIa=**100**;

end;

NHpopNoncMeanAll=estimate;

if Probt<**0.05** then NHpopNoncMeanSig=estimate;

if Probt<**0.05** and estimate>**0** then NHpopNoncMeanSigPos=estimate;

NHpopNoncTypeII=NHpopNoncTypeIIa+NHpopNoncTypeIIb;

/\*NHisClinMeanAll=.;

NHisClinMeanSig=.;

NHisClinMeanSigBene=.;

NHisClinType1a=0;

NHisClinType1b=0;

NHisClinTypeIa=0;

NHisClinTypeIb=0;

NHisClinTypeIIa=0;

NHisClinTypeIIb=0;

NHisClinType1=0;

NHisClinTypeI=0;

NHisClinTypeII=0;

NHisClinUnclear=0;

if Probt>0.05 then NHisClinUnclear=100;

NHisClinMeanAll=estimate;

if Probt<0.05 then do; \*interpret only sig effects;

if &ES>=&StdzedMagniThresh then do; \*true effect is beneficial; \*fix all these;

if -&StdzedMagniThresh<estimate<&StdzedMagniThresh then NHisClinTypeIIa=100;

if estimate<-&StdzedMagniThresh then NHisClinTypeIIb=100;

end;

if -&StdzedMagniThresh<&ES<&StdzedMagniThresh then do; \*true effect is trivial;

if estimate>&StdzedMagniThresh then NHisClinTypeIb=100;

if estimate<-&StdzedMagniThresh then NHisClinTypeIa=100;

end;

if &ES<=-&StdzedMagniThresh then do; \*true effect is harmful;

if estimate>&StdzedMagniThresh then NHisClinType1b=100;

if -&StdzedMagniThresh<estimate<&StdzedMagniThresh then NHisClinType1a=100;

end;

NHisClinMeanSig=estimate;

if estimate>&StdzedMagniThresh then NHisClinMeanSigBene=estimate;

NHisClinType1=NHisClinType1a+NHisClinType1b;

NHisClinTypeI=NHisClinTypeIa+NHisClinTypeIb;

NHisClinTypeII=NHisClinTypeIIa+NHisClinTypeIIb;

NHisClinTypeII=NHisClinTypeII+NHisClinType1;

end;

NHisNoncMeanAll=.;

NHisNoncMeanSig=.;

NHisNoncMeanSigPos=.;

NHisNoncTypeIIa=0;

NHisNoncTypeIIb=0;

NHisNoncTypeI=0;

NHisNoncTypeII=0;

NHisNoncUnclear=0;

if Probt>0.05 then NHisNoncUnclear=100;

NHisNoncMeanAll=estimate;

if Probt<0.05 then do; \*interpret only sig effects;

if &ES>=&StdzedMagniThresh then do; \*true effect is positive;

if -&StdzedMagniThresh<estimate<&StdzedMagniThresh then NHisNoncTypeIIa=100;

if estimate<-&StdzedMagniThresh then NHisNoncTypeIIb=100;

end;

if -&StdzedMagniThresh<&ES<&StdzedMagniThresh then do; \*true effect is trivial;

if estimate>&StdzedMagniThresh then NHisNoncTypeI=100;

if estimate<-&StdzedMagniThresh then NHisNoncTypeI=100;

end;

if &ES<=-&StdzedMagniThresh then do; \*true effect is negative;

if estimate>&StdzedMagniThresh then NHisNoncTypeIIb=100;

if -&StdzedMagniThresh<estimate<&StdzedMagniThresh then NHisNoncTypeIIa=100;

end;

NHisNoncMeanSig=estimate;

if estimate>0 then NHisNoncMeanSigPos=estimate;

NHisNoncTypeII=NHisNoncTypeIIa+NHisNoncTypeIIb;

end;

RetestCorr=&r;

SsizeA=&NoA;

SsizeB=&NoB;

StdzdEffect=&ES;\*/

**run**;

/\*KS: My corrected definitions of Type I and Type II error for clinical MBI\*/

/\*KS: Redefines unclear cases when the effect is beneficial as Type II errors, and also makes errors reflect a one-sided test\*/

**data** defsclinical;

set est1;

if estimate=**0** then do;

estimate=**.**; StdzedMagniThresh=**.**; EquivRawThresh=**.**; EquivPcentThresh=**.**; magni=""; clearornot="";

end;

rename df=DegFree;

CLpm=(Upper-lower)/**2**;

ClinTypeI=**0**;

ClinTypeII=**0**;

ClinUnclear=**0**;

if ES>=StdzedMagniThresh then do; \*true effect is beneficial;

if LCL99<-StdzedMagniThresh and UCL50>StdzedMagniThresh then ClinTypeII=**100**; \*KS: "UCLEAR" is the very definition of a type II error--you do a study and fail to find a real effect!;

\*if LCL99>-StdzedMagniThresh and UCL50>StdzedMagniThresh then correct; \*KS: at least possibly beneficial and implementable. Correct;

if -StdzedMagniThresh<LCL99 and UCL50<StdzedMagniThresh then ClinTypeII=**100**; \*KS: trivial is a type II error;

if -StdzedMagniThresh>LCL99 and UCL50<StdzedMagniThresh then ClinTypeII=**100**; \*KS: trivial to harmful is a Type II error;

if UCL90<-StdzedMagniThresh then ClinTypeII=**100**; \*KS: harmful is a type II error;

end;

if -StdzedMagniThresh<ES<StdzedMagniThresh then do; \*true effect is trivial;

\*if LCL99<-StdzedMagniThresh and UCL50>StdzedMagniThresh then correct; \*KS: "UCLEAR" is correct;

if LCL99>-StdzedMagniThresh and UCL50>StdzedMagniThresh then ClinTypeI=**100**; \*KS: Incorrectly conclude that the treatment is implementable=Type I error;

\*if -StdzedMagniThresh<LCL99 and UCL50<StdzedMagniThresh then correct; \*KS: (trivial) correctly conclude that the treatment is not beneficial;

\*if -StdzedMagniThresh>LCL99 and UCL50<StdzedMagniThresh then correct; \*KS: (trivial to harmful) correctly conclude that the treatment is not beneficial;

\*if UCL90<-StdzedMagniThresh then correct; \*KS: (harmful) correctly conclude that treatment is not beneficial;

end;

if ES<=-StdzedMagniThresh then do; \*true effect is harmful;

\*if LCL99<-StdzedMagniThresh and UCL50>StdzedMagniThresh then correct; \*KS: "UCLEAR" is correct since you did not declare a benefit;

if LCL99>-StdzedMagniThresh and UCL50>StdzedMagniThresh then ClinTypeI=**100**; \*KS: Conclude beneficial when you shouldn't. This is a Type I error.

Hopkins/Batterham have put a stricter guard against this Type I error ;

\*if -StdzedMagniThresh<LCL99 and UCL50<StdzedMagniThresh then correct; \*KS: (trivial) Not beneficial is correct.;

\*if -StdzedMagniThresh>LCL99 and UCL50<StdzedMagniThresh then correct; \*KS: (trivial to harmful). Not beneficial is correct;

\*if UCL90<-StdzedMagniThresh then correct; \*KS: (harmful). Not beneficial is correct;

end;

/\*KS: We have to compare to a one-sided standard hypothesis test\*/

NHpopClinTypeI=**0**; \*One-sided hypothesis test for benefit;

NHpopClinTypeII=**0**;

if ES>=StdzedMagniThresh then do; \*true effect is beneficial;

if Probt>**0.10** or estimate<**0** then NHpopClinTypeII=**100**; \*KS: Failing to find a positive effect is a type II error;

end;

if ES<StdzedMagniThresh then do; \*true effect is not beneficial;

if Probt<**0.10** and estimate>**0** then NHpopClinTypeI=**100**; \*KS: finding a positive effect is type I;

end;

where clinflag=**1**; \*Just reduces the double listing of trials. Calculations not dependent on this parameter;

**run**;

/\*KS: Corrected definitions for non-clinical MBI.\*/

/\*KS: Counts unclear cases correctly\*/

/\*KS: Counts inferences in the wrong direction when the effect is non-trivial as Type III errors.

/\*KS: Allows partial Type I and Type II errors, reflecting inference categories that Hopkins and Batterham propose.

/\*KS: I parsed the trivial-to-positive category into three inferences: Unlikely positive is a 15% type I error/85% type II error.

Possibly positive is 50/50.

Likely positive is 85% Type I/15% Type II.

These are based on the midpoints of the probability ranges provided by Hopkins/Batterham.

If you change these, it simply has the effect of making slightly different tradeoffs between Type II and Type I error\*/

**data** defsnonc;

set est1;

if estimate=**0** then do;

estimate=**.**; StdzedMagniThresh=**.**; EquivRawThresh=**.**; EquivPcentThresh=**.**; magni=""; clearornot="";

end;

rename df=DegFree;

CLpm=(Upper-lower)/**2**;

NoncTypeI=**0**;

NoncTypeII=**0**;

NoncTypeIII=**0**;

if ES>=StdzedMagniThresh then do; \*true effect is positive;

if LCL90<-StdzedMagniThresh and UCL90>StdzedMagniThresh then NoncTypeII=**100**; \*KS: Uclear effects are type II errors when there's a true meaningful effect, because you did a study and missed it!;

\*if LCL90>StdzedMagniThresh then correct; \*KS: positive effect is correct;

if LCL90>-StdzedMagniThresh and UCL90>StdzedMagniThresh and LCL90<StdzedMagniThresh then do; \*KS: trivial to positive categories;

if LCL50>StdzedMagniThresh then NoncTypeII=**15**; \*KS: likely positive category;

if LCL50<StdzedMagniThresh and UCL50>StdzedMagniThresh then NoncTypeII=**50**; \*KS: possibly positive;

if UCL50<StdzedMagniThresh then NoncTypeII=**85**; \*KS:unlikely positive 5-25%;

end;

if -StdzedMagniThresh<LCL90 and UCL90<StdzedMagniThresh then NoncTypeII=**100**; \*KS:Choosing trivial is a full type II error;

if LCL90<-StdzedMagniThresh and UCL90<StdzedMagniThresh and UCL90>-StdzedMagniThresh then do; \*KS: trivial to negative categories;

if UCL50<-StdzedMagniThresh then NoncTypeIII=**85**; \*KS: likely negative case. Type III here;

if UCL50>-StdzedMagniThresh and LCL50<-StdzedMagniThresh then NoncTypeIII=**50**; \*KS: possibly negative case;

if LCL50>-StdzedMagniThresh then NoncTypeIII=**15**;\*KS:unlikely negative 5-25%;

end;

if UCL90<-StdzedMagniThresh then NoncTypeIII=**100**; \*KS:declaring it negative is a full type III error;

end;

if -StdzedMagniThresh<ES<StdzedMagniThresh then do; \*true effect is trivial;

\*if LCL90<-StdzedMagniThresh and UCL90>StdzedMagniThresh then correct; \*KS: unclear is correct;

if LCL90>StdzedMagniThresh then NoncTypeI=**100**; \*KS: positive effect is type I error;

if LCL90>-StdzedMagniThresh and UCL90>StdzedMagniThresh and LCL90<StdzedMagniThresh then do;

if LCL50>StdzedMagniThresh then NoncTypeI=**85**; \*KS: likely positive category;

if LCL50<StdzedMagniThresh and UCL50>StdzedMagniThresh then NoncTypeI=**50**; \*KS: possibly positive;

if UCL50<StdzedMagniThresh then NoncTypeI=**15**; \*KS:unlikely positive 5-25%;

end;

\*if -StdzedMagniThresh<LCL90 and UCL90<StdzedMagniThresh then correct; \*KS:Choosing trivial correct;

if LCL90<-StdzedMagniThresh and UCL90<StdzedMagniThresh and UCL90>-StdzedMagniThresh then do; \*KS: trivial to negative categories;

if UCL50<-StdzedMagniThresh then NoncTypeI=**85**; \*KS: likely negative case. Type I here;

if UCL50>-StdzedMagniThresh and LCL50<-StdzedMagniThresh then NoncTypeI=**50**; \*KS: possibly negative case;

if LCL50>-StdzedMagniThresh then NoncTypeI=**15**;\*KS:unlikely negative 5-25%;

end;

if UCL90<-StdzedMagniThresh then NoncTypeI=**100**; \*KS:declaring it negative is a full type I error;

end;

if ES<=-StdzedMagniThresh then do; \*true effect is negative;

if LCL90<-StdzedMagniThresh and UCL90>StdzedMagniThresh then NoncTypeII=**100**; \*KS: unclear is type II;

if LCL90>StdzedMagniThresh then NoncTypeIII=**100**; \*KS: positive effect is type III error;

if LCL90>-StdzedMagniThresh and UCL90>StdzedMagniThresh and LCL90<StdzedMagniThresh then do;

if LCL50>StdzedMagniThresh then NoncTypeIII=**85**; \*KS: likely positive category;

if LCL50<StdzedMagniThresh and UCL50>StdzedMagniThresh then NoncTypeIII=**50**; \*KS: possibly positive;

if UCL50<StdzedMagniThresh then NoncTypeIII=**15**; \*KS:unlikely positive 5-25%;

end;

if -StdzedMagniThresh<LCL90 and UCL90<StdzedMagniThresh then NoncTypeII=**100**; \*KS:Choosing trivial type II;

if LCL90<-StdzedMagniThresh and UCL90<StdzedMagniThresh and UCL90>-StdzedMagniThresh then do; \*KS: trivial to negative categories;

if UCL50<-StdzedMagniThresh then NoncTypeII=**15**; \*KS: likely negative case;

if UCL50>-StdzedMagniThresh and LCL50<-StdzedMagniThresh then NoncTypeII=**50**; \*KS: possibly negative case;

if LCL50>-StdzedMagniThresh then NoncTypeII=**85**;\*KS:unlikely negative 5-25%;

end;

\*if UCL90<-StdzedMagniThresh then correct; \*KS:declaring it negative is correct;

end;

NHpopNoncTypeI=**0**;

NHpopNoncTypeII=**0**;

NHpopNoncTypeIII=**0**;

if ES>=StdzedMagniThresh then do; \*true effect is beneficial;

if Probt>**0.05** then NHpopNoncTypeII=**100**; \*KS: Failing to find an effect is a type II error;

if Probt<**0.05** and estimate<**0** then NHpopNoncTypeIII=**100**; \*KS: Finding an effect in the wrong direction is type III error;

end;

if -StdzedMagniThresh<ES<StdzedMagniThresh then do; \*true effect is trivial;

if Probt<**0.05** and estimate>**0** then NHpopNoncTypeI=**100**; \*KS: finding any effect is type I;

if Probt<**0.05** and estimate<**0** then NHpopNoncTypeI=**100**;

end;

if ES<=-StdzedMagniThresh then do; \*true effect is harmful;

if Probt<**0.05** and estimate>**0** then NHpopNoncTypeIII=**100**; \*KS:This is a type III error ;

if Probt>**0.05** then NHpopNoncTypeII=**100**; \*KS:Type II error;

end;

where clinflag=**0**; \*just reduces the double listing of each trial. Calculations are the same regardless of ClinFlag;

**run**;

/\*KS: Example Graphs. This gives example code for the graphics made. This code is not exhaustive for all cases.\*/

**proc** **format**;

value hopk

**1**="Clinical MBI"

**0**="Standard HT"

**2**="Non-clinical MBI"

;

**run**;

goptions reset = all;

goptions htext=**3**;

axis1 label=(angle=**90**);

/\*Plot Type I error by sample size, where ES=trivial\*/

/\*This is using Hopkins' original definitions\*/

**proc** **means** data=defs1 mean; \*Type I error;

var ClinTypeI NHpopClinTypeI ;

output out=typeI mean= hopk trad;

by No StdzedMagniThresh ES ;

where clinflag=**1**;

**run**;

**data** typeI;

set typeI;

typeI=hopk; hopk=**1**; output;

typeI=trad; hopk=**0**; output;

**run**;

**proc** **gplot** data=typeI;

plot typeI\*no=hopk/vaxis=axis1;

symbol1 v=dot c=blue i=join;

symbol2 v=dot c=red i=join;

symbol3 v=dot c=green i=join;

label No="Sample size per group";

label typeI="Type I error (%)";

label hopk="Method:";

where ES=**0** ;

title "Effect size=0";

Format hopk hopk.;

**run**;

**proc** **gplot** data=typeI;

plot typeI\*no=hopk/vaxis=axis1 ;

symbol1 v=dot c=blue i=join;

symbol2 v=dot c=red i=join;

symbol3 v=dot c=green i=join;

label No="Sample size per group";

label typeI="Type I error (%)";

label hopk="Method:";

where ES=**0.1** ;

title "Effect size=0.1";

Format hopk hopk.;

**run**;

**proc** **means** data=defsclinical mean; \*Correct clinical definitions;

var ClintypeI NHpopClinTypeI ;

output out=typeI mean= hopk trad ;

by No StdzedMagniThresh ES ;

where clinflag=**1**;

**run**;

**data** typeI;

set typeI;

typeI=hopk; hopk=**1**; output;

typeI=trad; hopk=**0**; output;

**run**;

**proc** **gplot** data=typeI;

plot typeI\*no=hopk/vaxis=axis1 ;

symbol1 v=dot c=blue i=join;

symbol2 v=dot c=red i=join;

symbol3 v=dot c=green i=join;

label No="Sample size per group";

label typeI="Type I error (%)";

label hopk="Method:";

where ES =0 ;

title "Effect size=0";

Format hopk hopk.;

**run**;

**run**;

**proc** **gplot** data=typeI;

plot typeI\*no=hopk/vaxis=axis1 ;

symbol1 v=dot c=blue i=join;

symbol2 v=dot c=red i=join;

symbol3 v=dot c=green i=join;

label No="Sample size per group";

label typeI="Type I error (%)";

label hopk="Method:";

where ES ne 0.1;

title "Effect size=0.1";

Format hopk hopk.;

**run**;

**proc** **means** data=defsnonc mean; \*Correct nonclinical definitions;

var nonctypeI NHpopNoncTypeI ;

output out=typeI mean= hopk trad ;

by No StdzedMagniThresh ES ;

where clinflag=**0**;

**run**;

**data** typeI;

set typeI;

typeI=hopk; hopk=**2**; output;

typeI=trad; hopk=**0**; output;

**run**;

**proc** **gplot** data=typeI;

plot typeI\*no=hopk/vaxis=axis1 ;

symbol1 v=dot c=blue i=join;

symbol2 v=dot c=green i=join;

symbol3 v=dot c=green i=join;

label No="Sample size per group";

label typeI="Type I error (%)";

label hopk="Method:";

where ES=0 ;

title "Effect size=0";

Format hopk hopk.;

**run**;

**proc** **gplot** data=typeI;

plot typeI\*no=hopk/vaxis=axis1 ;

symbol1 v=dot c=blue i=join;

symbol2 v=dot c=green i=join;

symbol3 v=dot c=green i=join;

label No="Sample size per group";

label typeI="Type I error (%)";

label hopk="Method:";

where ES=**0.1** ;

title "Effect size=0";

Format hopk hopk.;

**run**;

/\*KS: Additional code from Hopkins/Batterham that I did not use was truncated from this document\*/