User Guide: Macros for Patient-Specific Meta-Analysis (PSMA) Using Logistic Regression

14 September 2022

The patient-specific meta-analysis (PSMA) macros implement methods described in Crager and Tang (2014). Briefly, PSMA provides a method for estimating risk using meta-analysis methods applied to several studies. The PSMA calculation optimizes the use of the risk information provided by each study for patients like the one for whom the risk is being assessed based on their covariate values.

For a binary outcome, the risk (probability) of a specified outcome for a patient with a specified combination of covariate values is estimated using a weighted average of the log odds estimates from each study, weighting each estimate by the inverse of its variance. This calculation may be done using either fixed or random study effects. The weighted average log odds and the associated confidence interval endpoints are then transformed to the risk scale.

Under certain assumptions, PSMA may be applied where some the studies have special subpopulations not included in the other studies. The required assumptions are that (1) there is no interaction between subpopulation and the covariates used to estimate risk, and (2) if the subpopulations had been included in the other studies, subpopulation membership would have had a similar effect on risk as it did in the studies that included the subpopulation (Crager and Tang 2014). Assumption (1) can be tested using the data in the analysis but assumption (2) must be argued theoretically.

The risk estimates can be calculated for any future patient using a set of key summary statistics calculated once and for all for each study. The key summary statistics for proportional hazards regression are calculated using the key_stats_logistic macro. The macro is called as follows:

```
%key_stats_logistic(
    /* Input Specification */  indsn=,response=,event=,vars=,weight=,sampstrata=,
    /* Output Specification */ keysdsn=
    );
```

The macro parameters are described in Table 1.

Parameter	Type	Required?	Default Value	Description
indsn	$	Yes	(at temporary library)	(Libname reference and) file name containing input data set.
response	#/$	Yes	—	The dichotomous dependent variable for the logistic regression analysis.
event	#/$	Yes	—	The value of the response variable that indicates presence of an event.
vars	#	Yes	—	List of input data set variables containing the logistic regression model covariates.
weight	#	No	—	Input data set variable giving the observation's weight in the analysis. If this parameter is set, it is assumed that cohort sampling was used and resulted in the specified weights.
sampstrata	#/$	No	—	If a stratified cohort sampling design was used, use this parameter to list the stratification variables.
keydsn	$	Yes	—	(Libname reference and) output data set name.

Table 1. Macro key_stats_logistic Parameters

After the key summary statistics are computed for each study, risk estimates using PSMA with fixed study effects are calculated using macro risk_est_PSMA_logistic or with random study effects using macro risk_est_PSMA_logistic. Macro risk_est_PSMA_logistic is called as follows:

```
%risk_est_PSMA_logistic(
  /* Input parameters */   covdsn=,common_vars=,num_studies=,
                 special_pop_inds_1=,special_pop_inds_2=,special_pop_inds_3=,
                 special_pop_inds_4=,special_pop_inds_5=,special_pop_inds_6=,
                 special_pop_inds_7=,special_pop_inds_8=,special_pop_inds_9=,
                 key_stats1=,key_stats2=,key_stats3=,
                 key_stats4=,key_stats5=,key_stats6=,
                 key_stats7=,key_stats8=,key_stats9=,
  /* Analysis settings */  alpha=,
  /* Output parameters */  outdsn=,risk_est=,risk_LCL=,risk_UCL=
```

);

The macro parameters are described in Table 2.

Table 2. Macro risk_est_PSMA_logistic Parameters				
Parameter	Type	Required?	Default Value	Description
covdsn	$	Yes	(at temporary library)	(Libname reference and) name of the file containing covariate values at which to estimate the event risk. This data set must contain the variables specified in the parameters common_vars and special_pop_ind's. The input data set may have multiple rows.
common_vars	#	Yes	—	List of covariates common to all studies that are used to estimate the risk of the event.
num_studies	#	Yes	—	Constant giving the number of studies in the meta-analysis.
special_pop_inds_ <k>, <k>=1, 2, …, &num_studies.	#	No	—	Names of indicator functions for special populations, if any, in study <k>.
Key_stats<k>, <k> = 1, 2, …, &num_studies.	#	No	—	Input data set giving key summary statistics for study <k>. This data set must contain the following variables: beta_<var>, the Cox regression parameter estimate for each covariate <var> v_<var1>_<var2>, the covariance of the regression parameter estimates for each pair of covariates <var1> and <var2>.
alpha	#	No	0.05	Type I error rate to be used for calculating confidence intervals.
outdsn	$	Yes	—	Name of the data set that will contain the risk estimates.
Riskest	#	No	Riskest	Name of the output data set variable that will contain the risk estimates.
Risk_LCL	#	No	Risk_LCL	Name of the output data set variable that will contain the lower limit of the confidence interval for the risk.
Risk_UCL	#	No	Risk_UCL	Name of the output data set variable that will contain the lower limit of the confidence interval for the risk.

Macro risk_est_PSMA_logistic_rand is called as follows:

%risk_est_PSMA_logistic_rand(

  /* Input parameters */   covdsn=,common_vars=,num_studies=,

               special_pop_inds_1=,special_pop_inds_2=,special_pop_inds_3=,

               special_pop_inds_4=,special_pop_inds_5=,special_pop_inds_6=,

               special_pop_inds_7=,special_pop_inds_8=,special_pop_inds_9=,

               key_stats1=,key_stats2=,key_stats3=,

               key_stats4=,key_stats5=,key_stats6=,

               key_stats7=,key_stats8=,key_stats9=,

  /* Analysis settings */ convergence=,alpha=,

  /* Output parameters */ outdsn=,risk_est=,risk_LCL=,risk_UCL=

               );

The macro parameters are described in Table 3.

Table 3. Macro risk_est_PSMA_logistic_rand Parameters

Parameter	Type	Required?	Default Value	Description
covdsn	$	Yes	(at temporary library)	(Libname reference and) name of the file containing covariate values at which to estimate the event risk. This data set must contain the variables specified in the parameters common_vars and special_pop_ind's. The input data set may have multiple rows.
common_vars	#	Yes	—	List of covariates common to all studies that are used to estimate the risk of the event.
strata	$/#	No	—	If stratification was used in calculating the key statistics, list the stratification variables here.
num_studies	#	Yes	—	Constant giving the number of studies in the meta-analysis.
Special_pop_inds _<k>, <k>=1, 2, …, &num_studies.	#	No	—	Names of indicator functions for special populations, if any, in study <k>.

Table 3. Macro risk_est_PSMA_logistic_rand Parameters

Parameter	Type	Required?	Default Value	Description
Key_stats<k>, <k> = 1, 2, …, &num_studies.	#	No	—	Input data set giving key summary statistics for study <k>. This data set must contain the following variables: beta_<var>, the Cox regression parameter estimate for each covariate <var> v_<var1>_<var2>, the covariance of the regression parameter estimates for each pair of covariates <var1> and <var2>
Convergence	#	No	0.00000001	Convergence criterion for Newton-Raphson iteration to estimate interstudy variance.
alpha	#	No	0.05	Type I error rate to be used for calculating confidence intervals.
outdsn	$	Yes	—	Name of the data set that will contain the risk estimates.
Riskest	#	No	Riskest	Name of the output data set variable that will contain the risk estimates.
Risk_LCL	#	No	Risk_LCL	Name of the output data set variable that will contain the lower limit of the confidence interval for the risk.
Risk_UCL	#	No	Risk_UCL	Name of the output data set variable that will contain the lower limit of the confidence interval for the risk.

References

Crager MR, Tang G (2014). Patient-specific meta-analysis for risk assessment using multivariate proportional hazards regression. *Journal of Applied Statistics* **41**:2676–2695. DOI: 10.1080/02664763.2014.925102.