U.S. FOOD & DRUG ADMINISTRATION

Example Statistical Analysis Plan for Bayesian Subgroup Analysis: Sharing of Information Across Subgroups

Overview

Clinical trial participants are heterogeneous in demographics, disease characteristics, biomarkers, or in any other potentially prognostic variable for a clinical outcome of interest or potentially predictive variable of the treatment effect on that outcome.

When evaluating drug treatments, determining how and to what extent a drug works in different participant subgroups can be addressed by statistical approaches that use results from every subgroup when understanding the treatment effect for a given subgroup. The sample estimates of the treatment effects will tend to vary more than the underlying treatment effects, because there are two components of variability in the collection of sample subgroup treatment effects – one, the variability in the underlying true effects and two, random sampling variability because we have samples and not the entirety of the subgroup within the target population. When many subgroups are evaluated, some sample estimated treatment effects will represent random highs and some will represent random lows.

Bayesian subgroup analysis with hierarchical models can be utilized to address these issues of random highs and random lows. Such models have a prior distribution placed on the subgroup treatment effects according to an exchangeability structure implemented with random effect distributions. The resulting posterior mean of a subgroup treatment effect borrows strength from all the data (data internal and data external to that subgroup), giving it increased precision relative to the subgroup's sample treatment effect estimate that uses only data within that subgroup. In a one-way structure, a subgroup treatment effect has a posterior mean that shrinks the sample estimate toward the overall estimated treatment effect by taking a weighted average of the two quantities. The weight used on the overall estimated treatment effect increases with decreased evidence of treatment effect heterogeneity, as measured by the variation between subgroups relative to the variation within the subgroups. Unlike the sample estimates, the shrinkage estimates of the subgroup treatment effects do not (in expectation) have more variation than the underlying treatment effects themselves and are more precise than the sample estimates that are based on subgroup only data.

Study Design and Statistical Analysis

Study XY-03 is a double-blind, parallel-group, two-treatment, randomized, controlled trial of drug vs. control. The primary analysis evaluates a time-to-event endpoint using a hazard ratio. The trial is multiregional with sites in Asia, Europe, North America, and South America. There is interest in pre-specified supplemental statistical analyses that seek to understand heterogeneity in treatment effects across regions and estimate treatment effects within regions.

Let μ_k $(k = 1, 2, 3, 4)$ denote the underlying log-hazard ratio of a time-to-event endpoint of interest across four regions: Asia, Europe, North America, and South America. Let $\hat{\delta}_k$ $(k = 1, 2, 3, 4)$ denote the estimated log-hazard ratios within the respective regions, based on using data only from the given region. We have the hierarchical model:

$$\hat{\delta}_k \sim N(\mu_k, \sigma_k^2) \qquad k = 1, 2, 3, 4$$

$$\sigma_1^2 = a_1;\ \sigma_2^2 = a_2;\ \sigma_3^2 = a_3;\ \sigma_4^2 = a_4,$$

where a_1, a_2, a_3 and a_4 are the standard error estimates

$$\mu_k \sim N(\mu,\ \tau^2) \qquad k = 1, 2, 3, 4 \text{ and}$$

$$\mu \sim N(0, 16),\ \ \tau \sim Half\text{-}Normal(1).$$

A variance of 16 for μ corresponds to placing 1/8 of an event at time zero for each arm when there is a one to one randomization, so it is a very weakly informative prior. According to Spiegelhalter, et. al., Röver, et. al. and Neuenschwander, et. al., a Half-Normal(1) distribution for τ is a very conservative choice for modeling heterogeneity in logarithmic transformed statistics – such as a log-hazard ratio or log-odds ratio. The above model links the underlying within-region log-hazard ratios with each other. The outcomes from all participants are relevant in estimating the treatment effect within a given region with an outcome from any participant in that given region being more relevant than the outcome from any participant outside that region.

Software

```
 /* SAS code */
/*************************************************************************************
********\
 Summary level Region shrinkage analysis
 Sample estimate and lower and upper confidence limit of treatment effect in each subgroup
 1 - Asia
 2 - Europe
 3 - North America
 4 – South America
\*************************************************************************************
********/
data Region;
input subgroup $ loghr s2;
datalines;
/* Replace the loghr's and se's with the corresponding observed values*/
1 loghr1 se1
2 loghr2 se2
3 loghr3 se3
4 loghr4 se4
;
run;
  proc mcmc data=Region outpost=nlout seed=432 nmc=500000 thin=10 monitor=(HR)
STATISTICS=(summary interval);
  array HR[4];
  parms  mu tau2;
  prior mu ~ normal(0, var=16);
  prior tau2~ normal(0,sd=1,lower=1e-12);
  random theta ~n(mu, sd=tau2) subject=subgroup;
  HR[subgroup]=exp(theta);
```

```
  model loghr ~ n(theta, var=s2);
  ods output PostSummaries = ps_est PostIntervals = ps_int;
  run;
  data temp; merge ps_est(keep=parameter mean stddev) ps_int(keep=parameter CredibleLower
CredibleUpper);
  by parameter;
  format subgroup $22. EstCI $50.;
  if parameter = "HR1" then subgroup="Region: Asia";
  else if parameter = "HR2" then subgroup="Region: Europe";
  else if parameter = "HR3" then subgroup="Region: North America";
  else if parameter = "HR4" then subgroup="Region: South America";
  EstCI = strip(put(mean,8.2))||' ('||strip(put(CredibleLower,8.2))||', '||strip(put(CredibleUpper,8.2))||')';
  run;

  proc print data=temp;
  run;
```

References

Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian approaches to clinical trials and healthcare evaluation. John Wiley & Sons, Ltd; 2004: 67.

Röver C, Bender R, Dias S, Schmid CH, Schmidli H, Sturtz S, Weber S, Friede T. On weakly informative prior distributions for the heterogeneity parameter in Bayesian random-effects meta-analysis. Res Synth Methods. 2021;12(4):448-474. doi:10.1002/jrsm.1475.

Neuenschwander B, Weber S, Schmidli H, O'Hagan A. Predictively Consistent Prior Effective Sample Sizes. Biometrics, 2020(76), 578-587.

Henderson, N.C., Louis, T.A., Wang, C. and Varadhan, R. Bayesian analysis of heterogeneous treatment effects for patient-centered outcomes research. Health Services and Outcomes Research Methodology. 2016, 16(4), 213--233. https://doi.org/10.1007/s10742-016-0159-3 .

FDA Impact Story: Using innovative statistical approaches to provide the most reliable treatment outcomes information to patients and clinicians. Accessed May 8, 2024. https://www.fda.gov/drugs/regulatory-science-action/impact-story-using-innovative-statistical-approaches-provide-most-reliable-treatment-outcomes

Gelman, A. Prior distributions for variance parameters in hierarchical models. Bayesian Anal. 2006, 1, 515—534.

Pennello, G, Rothmann, MD (2018). Bayesian Subgroup Analysis with Hierarchical Models. In: Peace, K., Chen, DG., Menon, S. (eds) Biopharmaceutical Applied Statistics Symposium. ICSA Book Series in Statistics. Springer, Singapore. https://doi.org/10.1007/978-981-10-7826-2_10