Discussion: Recent Advances for Handling Composite Endpoints in Clinical Trials

Atalanta Ghosh, Surya Mohanty, Janssen Research and Development

Common Theme of the session: Zaslavsky talks about Bayesian approach of combining two or more endpoints into a composite and describes calculating weights for combining endpoints. Use of importance of endpoints into composite via γ_{L12} or p_{L12} was also discussed. Lachin introduced Wei-Lachin test for one-directional composite constructed using coefficients of individual endpoints. The test can be constructed using simple un-weighted average of model coefficients. The test can be modified to weighted test using weights according to clinical importance. The test performs better vs time-to-first composite event. Luo et al., introduces composite endpoints via win ratio with ordered outcomes. They described weighted win ratio as a better alternative for constructing certain composite endpoints. Importance of endpoints is introduced through natural and/or clinical ordering of endpoints. They also described utility of composite endpoint vs complication of construction and interpretation of composites. Bebu describes combination of endpoints using win ratio and proportion in favor concept. Importance of endpoints is incorporated as well in the construct ion of such composites. He showed the use of U-statistics for making inference based on these composite endpoints. Also possible extensions to 3 or more treatment comparisons are presented.

Bayesian Approach to Design and Analysis of Composite Endpoints in Clinical Trials with Multiple Dependent Binary Outcomes: This talk describes the composite endpoints for binary type endpoints. The composite is defined as Σ(weighted multiple endpoints). The endpoints can be utilized using equal weights. Unequal weights are proposed for these composites where weights are proportional to rates of events or clinical importance or both. Inference is based on multinomial distributions for proportions and uses Dirichlet conjugate prior resulting in Dirichlet posterior.

The composite of L_1 and L_2 disjointed sets of endpoints can be defined as $\Theta_L = \Theta_{L1} + C_{L12}*\Theta_{L2}$. The coefficient C_{L12} can be obtained for a given posterior pə, $P(\Theta_{L1} \leq C_{L12}*\Theta_{L2} |data) = p$. Conversely, posterior p can be obtained for a given non-inferiority margin C_{L12}. The weighted coefficient for L_1 and L_2 space was proposed such that, $P(\Theta_{L1}/C_{L12} \leq \Theta_{L2}) = p_{L12}$. The value of $p_{L12} = 0.5$ implies equal weights. Larger p_{L12} means less importance of L_1 endpoints. Whereas, $p_{L12} = 1$ implies no relevance of endpoints in L_1 space for the composite. Clinical importance can be derived via γ_{L12} $[=2(1-p_{L12})]$ relating contribution of L_1 vs L_2 $(0.5 <= p_{L12} <= 1)$.

A Multivariate One-Directional Test of Multiple Event-time Outcomes: Wei-Lachin analysis takes first event of each endpoint for each subject. The MACE does not capture the total impact of the disease status. Wei-Lachin test was a multivariate linear rank test with 1 df test based on the sum of the ranks for joint null against "stochastic ordering" alternative. The multivariate test is a one-sided or one-directional test where the implicit assumption is that the endpoints are all moving in the same direction. The individual endpoints in the composite cannot have different directions. The test is based on the unweighted mean of the treatment group coefficients from individual Cox proportional hazards models $Z=(\bar{\beta}_a + \bar{\beta}_b)/\bar{\sigma}$ or simply the average of the coefficients of the individual endpoints (2 or more as in PEACE study). A weighted test

can also be defined as $Z=(w_1*\bar{\beta}_a + w_2*\bar{\beta}_b)/\bar{\sigma}_w$, where $w_1+w_2=1$. Weights can be chosen to reflect clinical importance of the endpoints. This test is usually more powerful. As in the tests presented in the presentations in this forum, this test performs better compared to time-to-first composite test.

Weighted Win Ratio Approach:
This talk starts off with providing options for combining endpoints. The endpoints can be tested each using multiple testing with all endpoints or combining in to a composite and test the composite. Composite endpoint is useful for increasing event rate and/or effect size, reduction of study length, and sample size. These composites often times difficult to construct and interpret. The authors illustrated that there are two ways to combine the endpoints in a testing situation. The first is the traditional method where one combines the endpoints for single subject first then obtains the composite and performs treatment comparison, such as minimum or mean of endpoints. The second and a better method is to compare pairwise between treatment and control, calculate # wins and # loses and then construct either Win ratio (such as Ψ as introduced in the latter presentation) or the difference between them Δ = wins – loses. They discuss prioritizing endpoints using (1) Time to first: in this setting all other info (other than the first) is lost, (2) Average response: the responses (y_k) can be weighted $[\Sigma(w_k*y_k)/\Sigma w_k]$. The natural question in this case to choosing the weights, w_k. The win ratio approach prioritizes endpoints by comparing important ones first and naturally ordering the endpoints. Weighted win ratio in the context of time to event data defines winners under different scenarios. The authors discusses win ratio and first event analysis and weighing of win ratio by categorizing (W_2,L_1) and (L_2,W_1) pairs in their example. Possible weights can be (1) log-rank weights for terminal events, (2) at-risk probability of terminal events/non-terminal/or both. The optimal alternative hypothesis can be derived for given weights and optimal weights can be obtained for given alternative hypothesis space. They showed that weighted win ratio can provide improvement over non-weighted win ratio through example.

Large Sample Inference for a Win Ratio Analysis of a Composite Outcome:
Very nicely describes through examples disadvantages of traditional time to first composite event approach. Clearly defines win ratio and proportion in favor concept through visual example. It can be seen that Time to first event ~exp($\lambda d + \lambda s$) when the individuals follow exponential distribution. One can then have situation where ([$\lambda d - \delta$] + [$\lambda s + \delta$]) implying benefit in death endpoint (λd) but worsening in the stroke endpoint (λs) giving the same outcome for the composite, time to first analysis. In a similar manner one can realize the same composite outcome ($\lambda d + \lambda s$) with individual components as ([$\lambda d + \delta$] + [$\lambda s - \delta$]) where worsening in death but improving in stroke endpoint. Clearly, the first outcome is always preferable (with same final result) but the importance of endpoint is lost. It is important to know if a new therapy performs better on the most severe endpoint compared to standard of care. The author defined Ψ=win ratio and Δ=proportion in favor allow partial ordering in to account. He clearly describes the concept with paired (new vs standard) approach, where Ψ=# of winners/# of losers = τ_1/τ_2 and $\Delta = \tau_1 - \tau_2$. If U_1=# of winners for 1st grp/N and U_2=# of winners for 2nd grp/N, then one can use $U_1 - U_2$ for Δ and $R=U_1/U_2$ for Ψ. It can be shown that $\sqrt{N}(U_1-U_2) \sim N(\Delta, \sigma_\Delta)$ and $R \sim N(\Psi, \sigma_\Psi)$. One can use these for hypothesis testing and confidence intervals estimation. Extension to 3 or more comparisons is also be possible.

Other Approaches:
The topic of constructing composite endpoint has been around for quite some time. There

have been other approaches used in the past and some of them are listed here. One such attempt to construct composite endpoint is Time-to-event analysis of the endpoints in a given study. Other parametric and semi-parametric approaches were also attempted, such as Cox proportional hazard based models, Wei-Lin-Weissfeld (WLW) marginal method, Lee-Wei-Amato (LWA) overall effect method, Prentice-Williams-Peterson (PWP) conditional method and Anderson-Gill (AG) counting process method.

A remark on Clinical Importance:

It is important to consider clinical importance in construction of composite endpoints. Clinicians may value one endpoint over other differently than patients may view of importance. Incorporation of clinical importance is normally accomplished via the clinician's (single or a group) choice. Sometimes importance can be derived by patients' choice or preference. A natural question is how one can quantify such importance in a given setting. A Bayesian or other modeling approach can be proposed for quantification of clinical importance in construction of composite endpoints.

Papers (presentations) reviewed in this session (#672):

Bayesian Approach to Design and Analysis of Composite Endpoints in Clinical Trials with Multiple Dependent Binary Outcomes. Boris G. Zaslavsky, FDA/CBER

A Multivariate One-Directional Test of Multiple Event-Time Outcomes. John M. Lachin and Ionut Bebu, The George Washington University

Weighted Win Ratio Approach. Junshan Qiu and Steven Bai and Wei-Yann Tsai and Xiaodong Luo, FDA and FDA and Columbia University and Sanofi

Large Sample Inference for a Win Ratio Analysis of a Composite Outcome. Ionut Bebu* and John M. Lachin, The George Washington University