Discussion to: Choosing Monitoring Boundaries: Balancing Risks and Benefits

John M. Lachin

Research Professor of Biostatistics and Epidemiology, and of Statistics
The Biostatistics Center
Department of Epidemiology and Biostatistics
The George Washington University
jml@bsc.gwu.edu

April 19, 2017

What is the Question?

Is there a difference between the treatment groups in the set of primary outcomes?

- A test of significance

What is the nature of the difference between groups in ths set of outcomes?

- Parameter estimate and confidence limits.
- Often using a summary statistic (e.g. Win-Ratio)

Issues:

- What is the power and robustness of the test of a difference
- What is the clinical utility of the description of the difference(s)

Wei-Lachin Test

Wei and Lachin (*JASA*, 1984) describe a multivariate linear rank test for $K \geq 2$ measures.

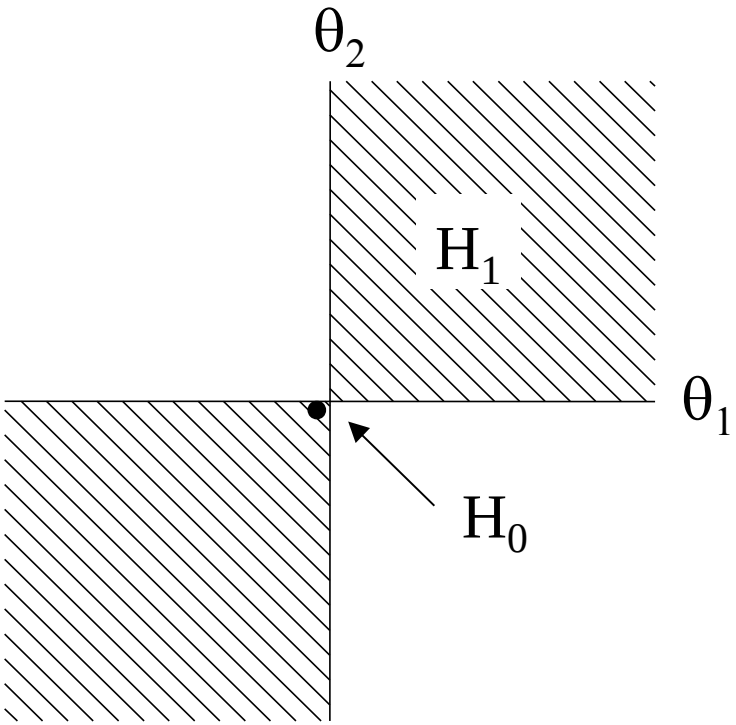
LJ Wei proposed a simple 1 *df* test of "stochastic ordering" that is a test of the joint null H_0 versus a multivariate one-directional (one-sided) alternative hypothesis.

Frick (*Commun. Statist.*, 1994) shows that the test is maximin efficient relative to the optimal (but unknown) test for the true (but unknown) parameters.

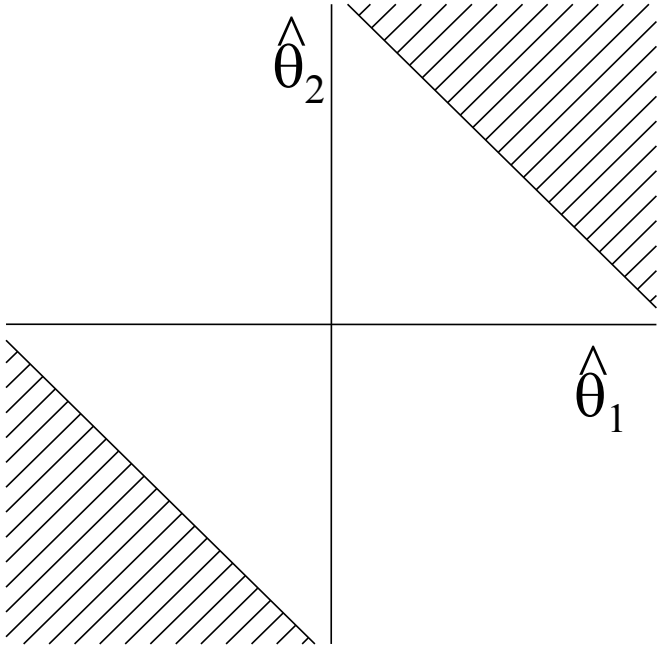
Lachin (*PLoS ONE*, 2014) describes applications to multiple outcomes on possibly different scales.

Lachin and Bebu (*Clinical Trials*, 2015) describe applications to multiple event-times (e.g. MACE).

Alternative Parameter Space



Rejection Region



The Test

Group-specific estimates $\widehat{\mu}_{ij}$ with expectation μ_{ij}, $i = 1,2$; j $= a, b$.

$\widehat{\delta}_j$ is the group difference for jth outcome

Vector $\widehat{\Delta} = (\widehat{\delta}_a \ \widehat{\delta}_b)^T$ with expectation $\Delta = (\delta_a \ \delta_b)^T$.

With large samples

$$\widehat{\Delta} \sim \mathcal{N}(\Delta, \Sigma)$$

with covariance matrix Σ that is consistently estimable with elements

$$\Sigma = \begin{bmatrix} \sigma_a^2 = V(\widehat{\delta}_a) & \sigma_{ab} = Cov(\widehat{\delta}_a, \widehat{\delta}_b) \\ \sigma_{ab} & \sigma_b^2 = V(\widehat{\delta}_b) \end{bmatrix}.$$

The Test

The Wei-Lachin test is then provided by

$$Z_S = \frac{\mathbf{J}'\widehat{\Delta}}{\sqrt{\mathbf{J}'\widehat{\Sigma}\mathbf{J}}} = \frac{\widehat{\delta}_a + \widehat{\delta}_b}{\widehat{\sigma}_S}, \qquad \mathbf{J} = (1\ 1)'$$

$$\widehat{\sigma}_S^2 = \widehat{V}(\widehat{\delta}_a + \widehat{\delta}_b) = \left[\widehat{\sigma}_a^2 + \widehat{\sigma}_b^2 + 2\widehat{\sigma}_{ab}\right]$$

Asymptotically $Z_S \sim N(0,1)$ under H_0 from Slutsky's theorem.

The test rejects H_0 in favor of H_{1S} when $Z_S \geq Z_{1-\alpha}$ at level α one-sided.

A two-sided test would reject when $|Z_S| \geq Z_{1-\alpha/2}$.

Model Based Covariances

Model based estimates are readily obtained from partitioning the information sandwich estimates (Pipper, et al., *JRSS*, 2012).

The *mmm* function in the **R** package *multcomp*

Consider separate regression models for X_a and X_b.

Then the robust information sandwich estimate of the covariance matrix of the coefficients in each model are:

$$\begin{aligned}\mathbf{Cov}(\widehat{\theta}_a)_{K_a \times K_a} &= \mathbf{I}_a(\widehat{\theta}_a)^{-1}\mathrm{U}_a(\widehat{\theta}_a)\mathrm{U}_a(\widehat{\theta}_a)'\mathbf{I}_a(\widehat{\theta}_a)^{-1} \\ \mathbf{Cov}(\widehat{\theta}_b)_{K_b \times K_b} &= \mathbf{I}_b(\widehat{\theta}_b)^{-1}\mathrm{U}_b(\widehat{\theta}_b)\mathrm{U}_b(\widehat{\theta}_b)'\mathbf{I}_b(\widehat{\theta}_b)^{-1}.\end{aligned}$$

and the covariance is

$$\mathbf{Cov}(\widehat{\theta}_a, \widehat{\theta}_b)_{K_a \times K_b} = \mathbf{I}_a(\widehat{\theta}_a)^{-1}\mathrm{U}_a(\widehat{\theta}_a)\mathrm{U}_b(\widehat{\theta}_b)'\mathbf{I}_b(\widehat{\theta}_b)^{-1}.$$

Applies to multiple outcomes of different types with covariate adjustment.

Cardiovascular Outcome Trial Composite Analysis

A Wei-Lachin analysis would count the first of each type of event experienced by each patient.

Can increase power.

A composite time-to-first event outcome analysis does not capture the total disease burden.

May sacrifice power.

Multiple PH Models

$\beta_j = \log(HR)$ for the j-th outcome for E versus C.
$\beta_j < 0$ now favors E versus C.

The test then becomes

$$Z_S = \frac{\mathbf{J}'\widehat{\beta}}{\sqrt{\mathbf{J}'\widehat{\Sigma}\mathbf{J}}} = \frac{\widehat{\beta}_a + \widehat{\beta}_b}{\widehat{\sigma}_S} = \frac{\widehat{\overline{\beta}}}{\sqrt{\widehat{V}\left(\widehat{\overline{\beta}}\right)}}$$

Reject H_0 in favor of H_{1S} when $Z_S \leq Z_\alpha$.

Multiple PH Models, continued

Can also use a weighted combination of the estimates of the form

$$Z_{Sw} = \frac{\mathbf{W}'\widehat{\beta}}{\sqrt{\mathbf{W}'\widehat{\Sigma}\mathbf{W}}} = \frac{w_a\widehat{\beta}_a + w_b\widehat{\beta}_b}{\left[w_a^2\widehat{\sigma}_a^2 + w_b^2\widehat{\sigma}_b^2 + 2w_a w_b\widehat{\sigma}_{ab}\right]^{1/2}} = \frac{\widehat{\overline{\beta}}_w}{\sqrt{\widehat{V}\left(\widehat{\overline{\beta}}_w\right)}},$$

where $\mathbf{W}'\mathbf{J} = \mathbf{1}$, and $\mathbf{W}$ is pre-specified.

The weights can reflect the relative severity or importance of the component outcomes.

PEACE Study

The *Prevention of Events with Angiotensin Converting Enzyme Inhibition* (PEACE) study (*NEJM*, 2004)

Assessed whether treatment with an ACE inhibitor (ACEi , n=4158) versus placebo (n=4132) would reduce the risk of CVD

Consider the outcome MACE + CHF, or time to CVD death, non-fatal MI, non-fatal stroke, hospitalization for CHF

PEACE Outcomes

Numbers of subjects (cases) with each type of cardiovascular event and for the composite outcomes.

	# Cases				
	ACEi	Placebo	ACEi vs Placebo		One-sided
Outcome	(n=4158)	(n=4132)	HR	95% CI	p
CV death	146	152	0.95	0.76, 1.19	0.34
Non-fatal MI	222	220	1.0	0.83, 1.21	0.5
Non-fatal stroke	55	75	0.72	0.51, 1.03	0.035
CHF	105	134	0.77	0.6, 1.0	0.025
Composite	449	492	0.90	0.79, 1.02	0.06
Wei-Lachin					
One-sided	–	–	0.854	–, 0.964	0.016
Two-sided	–	–	–	0.74, 0.99	0.032

Other Methods Applied to the MACE + CHF analysis

MANOVA omnibus test

$\chi^2_4 = 7.39$ on 4-df with $p = 0.117$.

Weighted Wei-Lachin test with weights

Event:	CV Death	non-fatal MI	non-fatal stroke	non-fatal CHF
Weight:	0.5	0.1	0.25	0.15

with weights that sum to 1.0.

Analysis	HR	95% CI	p
Weighted Wei-Lachin	0.965	0.927, 1.003	0.037

Win-Ratio

Analysis	Ratio	95% CI	one-sided p
Win Ratio	1.11	0.973, 1.266	0.0941

Conclusions

The simple Wei-Lachin one-directional multivariate test is based on the sum of the component statistics, or the unweighted mean of the component model coefficients.

The test is maximin efficient when there is truly a preponderance of benefit for the set of outcomes, with no harm for any.

The test is more powerful than multiple tests with a multiplicity adjustment or a MANOVA omnibus test, when the one-directional multivariate hypothesis applies.

The test can be applied to mixtures of different variable types and can adjust for covariates.

For composite outcome event times, the test is largely superior to the common time-to-first-event composite analysis.

A note of Caution

The composite time to the first component event can be biased relative to the marginal analysis of the individual components.

Simulation using a shared frailty bivariate exponential model, equivalent to the Marshall-Olkin distribution.

Simulation Under Joint Marginal H_0

First consider a simulation under the joint null hypothesis H_0 where

$$\lambda_{1a} = \lambda_{2a} = \lambda_{1b} = \lambda_{2b} = 0.2,$$
$$\lambda_{1f} = 0.1, \text{ and}$$
$$\text{correlation } \rho_1 = 0.33$$

Then the properties for other values of the group 2 frailty and correlation ρ_2 are provided by

		No Censoring ($n = 100$)	With Censoring ($n = 200$)
λ_{2f}	ρ_2	α	α
0.100	0.333	0.0530	0.0474
0.075	0.231	0.0912	0.0716
0.050	0.143	0.1890	0.1342
0.025	0.067	0.3479	0.2366

Simulation Under Joint Marginal Alternative

Now assume

$$\lambda_{1a} = \lambda_{1b} = 0.3$$
$$\lambda_{2a} \neq \lambda_{2b} = 0.2$$
$$\lambda_{1f} \neq \lambda_{2f} = 0.10$$
$$\rho_2 \neq \rho_1 = 0.10 \text{ and}$$
$$\text{No censoring, } n = 100$$

Then the properties are provided by

				Prob. Reject	
λ_{2a}	λ_{1f}	ρ_1	ρ_2	Composite	Wei-Lachin
0.30	0.20	0.500	0.250	0.047	0.816
0.25	0.25	0.714	0.286	0.052	0.809

Simulation

Similar results also apply to the Win Ratio

However, even with different frailties (correlations) between groups, the following tests remain unaffected:

The 1 *df* Wei-Lachin test

Separate tests with a Bonferroni (Holm) adjustment.

A 2 *df* T^2-like omnibus or "MANOVA" test

Recommendation

Consider basing an inference on the *magnitude* of the difference between groups using a robust, efficient test such as the Wei-Lachin text

Then employ other summary measures to describe the *nature* of the group differences, such as the Win-ratio, recognizing that in general these approaches will be less powerful and some may be affected by unequal covariances.